Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Lauzon, Dominique (he/him, il)

Subject:	AI - Handover system requirements
Location:	Microsoft Teams Meeting
Start:	Wed 2023-08-23 1:00 PM
End:	Wed 2023-08-23 1:30 PM
Show Time As:	Tentative
Recurrence:	(none)
Meeting Status:	Not yet responded
Organizer:	Trew, Debbie
Required Attendees:	Cain, Martin (BDO Lixar); Thompson, Ross; Herbert, Wesley
Optional Attendees:	Lauzon, Dominique

Microsoft Teams meeting

Join on your computer, mobile app or room device
Click here to join the meeting

Meeting ID: 242 962 054 817
Passcode: dkGixu
Download Teams | Join on the web

Or call in (audio only)

+1 343-803-0870,,313455321# Canada, Ottawa-Hull

Phone Conference ID: 313 455 321#
Find a local number | Reset PIN

Learn More | Meeting options

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Lauzon, Dominique (he/him, il)

From: Thompson, Ross
Sent: September 20, 2023 9:56 AM
To: Vadnais, Philippe; Trew, Debbie; Annett, Megan
Cc: Mockett, Ryan; Sajjad, Rabia
Subject: RE: Search requirements
Attachments: og-ati-search_EN.json; og-ati-search_FR.json

UNCLASSIFIED / NON CLASSIFIÉ

Good morning Phlippe,

The link I previously sent you should contain complete facet information, but it is in a different format than you are used to. Fortunately I believe I can provide a response that it much closer to what Lixar is giving, and hopefully this will minimize the changes you need to make on your end.

For example, to make a query close to the example you provided, this URL will generate more familiar JSON:

https://search-staging.open.canada.ca/ati/?page=1&sort=score+desc&search_text=nrc&search_format=solr

The OG search response will contain more information and more facets then the one from Lixar, but hopefully because it is in JSON you can just ignore those fields that are irrelevant to you.

Regardless of the JSON response, there are a couple of differences of note with the OG search.

1. Search is always language specific in order to treat the languages equivalently. This means if you want English language results, you must use the hostname: **search-staging.open.canada.ca**, For French language results, use the host name: **stadification-rechercher.ouvert.canada.ca**. Except for this, the URLs are the same when using the search. For example:

   https://stadification-rechercher.ouvert.canada.ca/ai/?page=1&sort=score+desc&search_text=nrc&search_format=solr

2. To get the Solr style JSON, you must add search_format=solr to the URL, or the app will return a webpage.
3. To search for text, add search_text=< *your text* > to the URL.
4. Facet searches with multiple values are comma separated. Commas are URL encoded. For example to select both 2023 and 2023 years: https://search-staging.open.canada.ca/ati/?page=1&sort=score+desc&search_text=nrc&year=2023%7C2022&search_format=solr

In terms of the HTTP referrer-policy, this is what I am see in the HTTP response header when I call the URL. Let me know if this is not what you were looking for.

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Thanks,
Ross

IT Team Lead, Office of the Chief Information Officer
Treasury Board of Canada Secretariat / Government of Canada
Ross.Thompson@tbs-sct.gc.ca / Tel: 613.371.4187 / TTY: 613-369-9371

Chef d'équipe TI, Bureau du dirigeant principal de l'information
Secrétariat du Conseil du Trésor du Canada / Gouvernement du Canada
Ross.Thompson@tbs-sct.gc.ca / Tél : 613.371.4187 / ATS : 613-369-9371

From: Vadnais, Philippe
Sent: Friday, September 1, 2023 7:57 AM
To: Thompson, Ross ; Trew, Debbie ; Sajjad, Rabia
Cc: Mockett, Ryan
Subject: RE: Search requirements

UNCLASSIFIED / NON CLASSIFIÉ

Thank looking into it,

I should be clearer for the facet, Lixar right now return to the list of options we have on the search.

Let say you search for "car" in the json facet section you will have a list of institutions you can select with the count of request associates to the filter, basically all the information to build the facet structure on an html page. I joint an example of attachment. We don't need to follow exactly what Lixar was doing but at lease have the same info will help.

Let me know if you have any question.

Thank you and have a nice day,

From: Thompson, Ross <Ross.Thompson@tbs-sct.gc.ca>
Sent: Thursday, August 31, 2023 4:24 PM
To: Vadnais, Philippe <Philippe.Vadnais@tbs-sct.gc.ca>; Trew, Debbie <Debbie.Trew@tbs-sct.gc.ca>; Sajjad, Rabia <Rabia.Sajjad@tbs-sct.gc.ca>
Cc: Mockett, Ryan <Ryan.Mockett@tbs-sct.gc.ca>
Subject: RE: Search requirements

UNCLASSIFIED / NON CLASSIFIÉ

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Hi Philippe,

Regarding the two points:

1. So far as I know, the application should be returning a referrer-policy HTTP header of "unsafe-url". It appears to be that way when I browse it using Edge. Is it possible it is getting overwritten somewhere?
2. I have updated the search application to return a log more facet information

Best regards,
Ross

IT Team Lead, Office of the Chief Information Officer
Treasury Board of Canada Secretariat / Government of Canada
Ross.Thompson@tbs-sct.gc.ca / Tel: 613.371.4187 / TTY: 613-369-9371

Chef d'équipe TI, Bureau du dirigeant principal de l'information
Secrétariat du Conseil du Trésor du Canada / Gouvernement du Canada
Ross.Thompson@tbs-sct.gc.ca / Tél : 613.371.4187 / ATS : 613-369-9371

From: Vadnais, Philippe <Philippe.Vadnais@tbs-sct.gc.ca>
Sent: Thursday, August 31, 2023 8:32 AM
To: Thompson, Ross <Ross.Thompson@tbs-sct.gc.ca>; Trew, Debbie <Debbie.Trew@tbs-sct.gc.ca>; Sajjad, Rabia <Rabia.Sajjad@tbs-sct.gc.ca>
Cc: Mockett, Ryan <Ryan.Mockett@tbs-sct.gc.ca>
Subject: RE: Search requirements

UNCLASSIFIED / NON CLASSIFIÉ

Good morning everyone,

I was playing with the search via my own concepts html page.
I already identify 2 issues:

- The biggest is the security policy on the page to generate the json : Referrer Policy:strict-origin-when-cross-origin That means I am not able to call this url from JavaScript or the dotnet back end.
- The second issue is about the filtering with Lixar they were returning the facet for the filtering. Right now the json doesn't have those structure (it make sense people just want the info not the filter). This mean we will need to generate those filter on the client side that may impact the performance of the page.

Let me know if you have control on those security policy, I don't know how OG portal is configured or administered.

Thank you and have a nice day,

From: Thompson, Ross <Ross.Thompson@tbs-sct.gc.ca>
Sent: Wednesday, August 30, 2023 4:45 PM
To: Trew, Debbie <Debbie.Trew@tbs-sct.gc.ca>; Vadnais, Philippe <Philippe.Vadnais@tbs-sct.gc.ca>; Sajjad, Rabia <Rabia.Sajjad@tbs-sct.gc.ca>
Cc: Mockett, Ryan <Ryan.Mockett@tbs-sct.gc.ca>
Subject: RE: Search requirements

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Hi Debbie and Philippe,

Thanks for meeting this afternoon to discuss the ATI search replacement. Below are the links and some random technical details about searching.

The links to the draft version of the ATI search are:
(EN) https://search-staging.open.canada.ca/ati/
(FR) https://stadification-rechercher.ouvert.canada.ca/ai/

In the search, there are four fields that can be used to filter results:

1. disposition
2. owner_org
3. year
4. month

For example, to filter on records from TBS, that are "All disclosed", from 2022, the API call would be:

https://search-staging.open.canada.ca/ati/?page=1&[redacted]
&sort=year+desc&disposition=DA&[redacted]&search_format=json

You can turn any web page search, into a API call just by adding "&search_format=json" to the URL

To perform a general text search, filter using the search_text field. For example: https://stadification-rechercher.ouvert.canada.ca/ai/?search_text=contrat

The data from which this search is derived can be found on the Open Government portal at: https://open.canada.ca/data/en/dataset/0797e893-751e-4695-8229-a5066e4fe43c.

Right now, the search returns results either in English or in French. If we need to change the format, we can discuss next month.

Thanks,
Ross

IT Team Lead, Office of the Chief Information Officer
Treasury Board of Canada Secretariat / Government of Canada
Ross.Thompson@tbs-sct.gc.ca / Tel: 613.371.4187 / TTY: 613-369-9371

Chef d'équipe TI, Bureau du dirigeant principal de l'information
Secrétariat du Conseil du Trésor du Canada / Gouvernement du Canada
Ross.Thompson@tbs-sct.gc.ca / Tél : 613.371.4187 / ATS : 613-369-9371

-----Original Appointment-----
From: Trew, Debbie <Debbie.Trew@tbs-sct.gc.ca>
Sent: Wednesday, August 30, 2023 8:57 AM
To: Trew, Debbie; Vadnais, Philippe; Sajjad, Rabia; Thompson, Ross
Cc: Mockett, Ryan
Subject: Search requirements
When: Wednesday, August 30, 2023 1:30 PM-2:00 PM (UTC-05:00) Eastern Time (US & Canada).
Where: Microsoft Teams Meeting

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Hi
I thought it would be a good idea to discuss the ATIP Online replacing our current search with the OG Portal search.
What do we need to know in order to implement the change.

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
{
    "responseHeader": {
        "status": 0,
        "QTime": 10,
        "params": {
            "f.year.facet.sort": "index",
            "f.month.facet.limit": "12",
            "hl": "on",
            "f.owner_org.facet.limit": "200",
            "indent": "false",
            "fl":
"id,month,month_en,disposition,disposition_en,comments_en,umd_number,owner_org,owne
r_org_en,year,request_number,summary_en,pages,pages_en,owner_org_title",
            "f.umd_number.facet.sort": "count",
            "q.op": "AND",
            "f.owner_org.facet.sort": "label",
            "f.disposition.facet.limit": "50",
            "hl.simple.pre": "<mark>",
            "highlight": "true",
            "defType": "edismax",
            "qf": [
                "id",
                "month",
                "month_en",
                "disposition",
                "disposition_en",
                "comments_en",
                "umd_number",
                "owner_org",
                "owner_org_en",
                "year",
                "request_number",
                "summary_en",
                "pages",
                "pages_en",
                "owner_org_title"
            ],
            "sow": "true",
            "f.umd_number.facet.limit": "100",
            "hl.fl": [
                "disposition",
                "comments_en",
                "umd_number",
                "owner_org",
                "request_number",
                "summary_en",
                "owner_org_title"
            ],
            "f.disposition.facet.sort": "count",
            "wt": "json",
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
            "hl.highlightMultiTerm": "true",
            "facet.field": [
                "owner_org",
                "disposition",
                "year",
                "month",
                "umd_number",
                "owner_org_title"
            ],
            "start": "0",
            "f.owner_org_title.facet.sort": "count",
            "f.owner_org_title.facet.limit": "100",
            "sort": "score desc",
            "rows": "10",
            "hl.snippets": "10",
            "q": "nrc",
            "hl.simple.post": "</mark>",
            "f.month.facet.sort": "index",
            "f.year.facet.limit": "100",
            "facet.method": "enum",
            "hl.method": "unified",
            "facet.mincount": "1",
            "facet": "true",
            "facet.sort": "index"
        }
    },
    "response": {
        "numFound": 73,
        "start": 0,
        "numFoundExact": true,
        "docs": [
            {
                "year": 2022,
                "month": 6,
                "month_en": "June",
                "request_number": "A-2022-09",
                "summary_en": "Briefing note BN0007793 Update on the
<mark>NRC</mark>’s Clinical Trial Material Facility",
                "disposition": "DP",
                "disposition_en": "Disclosed in part",
                "pages": 3,
                "pages_en": "3",
                "comments_en": "-",
                "umd_number": "80",
                "owner_org": "nrc-cnrc",
                "owner_org_en": "National Research Council Canada",
                "owner_org_title": "National Research Council Canada | Conseil
national de recherches Canada",
                "id": "nrc-cnrc,2022,6,A-2022-09"
            },
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
            {
                "year": 2021,
                "month": 7,
                "month_en": "July",
                "request_number": "A-2021-27",
                "summary_en": "Briefing Note Ref # BN0007793 Update on the
<mark>NRC</mark>’s Clinical Trial Material Facility",
                "disposition": "DP",
                "disposition_en": "Disclosed in part",
                "pages": 3,
                "pages_en": "3",
                "comments_en": "-",
                "umd_number": "80",
                "owner_org": "nrc-cnrc",
                "owner_org_en": "National Research Council Canada",
                "owner_org_title": "National Research Council Canada | Conseil
national de recherches Canada",
                "id": "nrc-cnrc,2021,7,A-2021-27"
            },
            {
                "year": 2022,
                "month": 7,
                "month_en": "July",
                "request_number": "A-2022-14",
                "summary_en": "Summary of how many researchers use each of the
<mark>NRC</mark>’s networks (Black network, orange network and green network) as of
the March 18, 2022 cyber incident reported by the <mark>NRC</mark>. ",
                "disposition": "DA",
                "disposition_en": "All disclosed",
                "pages": 1,
                "pages_en": "1",
                "comments_en": "-",
                "umd_number": "80",
                "owner_org": "nrc-cnrc",
                "owner_org_en": "National Research Council Canada",
                "owner_org_title": "National Research Council Canada | Conseil
national de recherches Canada",
                "id": "nrc-cnrc,2022,7,A-2022-14"
            },
            {
                "year": 2022,
                "month": 8,
                "month_en": "August",
                "request_number": "A-2022-05",
                "summary_en": "All internal notes and communications at
<mark>NRC</mark>, as well as with other federal organizations, regarding the
establishment of the effective date of <mark>NRC</mark>'s mandatory vaccination
policy for November 8, 2021.",
                "disposition": "DP",
                "disposition_en": "Disclosed in part",
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
            "pages": 128,
            "pages_en": "128",
            "comments_en": "-",
            "umd_number": "80",
            "owner_org": "nrc-cnrc",
            "owner_org_en": "National Research Council Canada",
            "owner_org_title": "National Research Council Canada | Conseil
national de recherches Canada",
            "id": "nrc-cnrc,2022,8,A-2022-05"
        },
        {
            "year": 2022,
            "month": 2,
            "month_en": "February",
            "request_number": "A-2021-38",
            "summary_en": "<mark>NRC</mark>'s licensing agreement with CanSino
for the HK293 cell line (and any related contracts/agreements/extensions)",
            "disposition": "DP",
            "disposition_en": "Disclosed in part",
            "pages": 18,
            "pages_en": "18",
            "comments_en": "-",
            "umd_number": "80",
            "owner_org": "nrc-cnrc",
            "owner_org_en": "National Research Council Canada",
            "owner_org_title": "National Research Council Canada | Conseil
national de recherches Canada",
            "id": "nrc-cnrc,2022,2,A-2021-38"
        },
        {
            "year": 2021,
            "month": 8,
            "month_en": "August",
            "request_number": "A-2021-42",
            "summary_en": "Briefing note Ref # BN21-06-105, Approval of the
Evaluation of <mark>NRC</mark>'s Metrology Research Centre",
            "disposition": "DA",
            "disposition_en": "All disclosed",
            "pages": 54,
            "pages_en": "54",
            "comments_en": "-",
            "umd_number": "80",
            "owner_org": "nrc-cnrc",
            "owner_org_en": "National Research Council Canada",
            "owner_org_title": "National Research Council Canada | Conseil
national de recherches Canada",
            "id": "nrc-cnrc,2021,8,A-2021-42"
        },
        {
            "year": 2021,
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
                "month": 10,
                "month_en": "October",
                "request_number": "A-2021-01",
                "summary_en": "Documents related to the National Research Council
of Canada (<mark>NRC</mark>) decision to transfer ownership of equipment purchased
by <mark>NRC</mark> to Lanaupôle Fibers Inc. ",
                "disposition": "DP",
                "disposition_en": "Disclosed in part",
                "pages": 12,
                "pages_en": "12",
                "comments_en": "-",
                "umd_number": "80",
                "owner_org": "nrc-cnrc",
                "owner_org_en": "National Research Council Canada",
                "owner_org_title": "National Research Council Canada | Conseil
national de recherches Canada",
                "id": "nrc-cnrc,2021,10,A-2021-01"
            },
            {
                "year": 2021,
                "month": 12,
                "month_en": "December",
                "request_number": "A-2021-33",
                "summary_en": "Briefing note BN0008318 Support to Invoke National
Security Exception for Timely Design and Build of <mark>NRC</mark>'s Clinical Trial
Material Facility",
                "disposition": "DP",
                "disposition_en": "Disclosed in part",
                "pages": 7,
                "pages_en": "7",
                "comments_en": "-",
                "umd_number": "80",
                "owner_org": "nrc-cnrc",
                "owner_org_en": "National Research Council Canada",
                "owner_org_title": "National Research Council Canada | Conseil
national de recherches Canada",
                "id": "nrc-cnrc,2021,12,A-2021-33"
            },
            {
                "year": 2022,
                "month": 10,
                "month_en": "October",
                "request_number": "A-2022-08",
                "summary_en": "Briefing note BN0008318Support to Invoke National
Security Exception for Timely Design and Build of <mark>NRC</mark>'s Clinical Trial
Material Facility",
                "disposition": "DP",
                "disposition_en": "Disclosed in part",
                "pages": 7,
                "pages_en": "7",
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
                "comments_en": "-",
                "umd_number": "80",
                "owner_org": "nrc-cnrc",
                "owner_org_en": "National Research Council Canada",
                "owner_org_title": "National Research Council Canada | Conseil
national de recherches Canada",
                "id": "nrc-cnrc,2022,10,A-2022-08"
            },
            {
                "year": 2022,
                "month": 7,
                "month_en": "July",
                "request_number": "A-2022-12",
                "summary_en": "A copy of the email sent by <mark>NRC</mark>
President Iain Stewart to @<mark>NRC</mark> email addresses in March regarding the
cyber incident that affected the National Research Council on March 18, 2022. All
other subsequent emails from Iain Stewart/<mark>NRC</mark> managers or directors to
<mark>NRC</mark> staff/researchers/clients providing updates on the above cyber
incident since the original email. ",
                "disposition": "DP",
                "disposition_en": "Disclosed in part",
                "pages": 42,
                "pages_en": "42",
                "comments_en": "-",
                "umd_number": "80",
                "owner_org": "nrc-cnrc",
                "owner_org_en": "National Research Council Canada",
                "owner_org_title": "National Research Council Canada | Conseil
national de recherches Canada",
                "id": "nrc-cnrc,2022,7,A-2022-12"
            }
        ]
    },
    "facet_counts": {
        "facet_queries": {},
        "facet_fields": {
            "owner_org": [
                "aafc-aac",
                2,
                "cnsc-ccsn",
                5,
                "csa-asc",
                1,
                "csec-cstc",
                2,
                "dfo-mpo",
                1,
                "dnd-mdn",
                4,
                "ec",
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
    1,
    "fin",
    1,
    "hc-sc",
    2,
    "ic",
    3,
    "infc",
    2,
    "lac-bac",
    3,
    "nrc-cnrc",
    42,
    "nrcan-rncan",
    2,
    "statcan",
    1,
    "tbs-sct",
    1
],
"disposition": [
    "DP",
    47,
    "DA",
    13,
    "NE",
    11,
    "EC",
    1,
    "EX",
    1
],
"year": [
    "2012",
    2,
    "2013",
    2,
    "2014",
    4,
    "2015",
    2,
    "2016",
    1,
    "2017",
    1,
    "2018",
    2,
    "2019",
    4,
    "2020",
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
    13,
    "2021",
    17,
    "2022",
    22,
    "2023",
    3
],
"month": [
    "1",
    6,
    "2",
    9,
    "3",
    10,
    "4",
    1,
    "5",
    2,
    "6",
    5,
    "7",
    8,
    "8",
    12,
    "9",
    1,
    "10",
    4,
    "11",
    4,
    "12",
    11
],
"umd_number": [
    "80",
    42,
    "29",
    5,
    "32",
    4,
    "34",
    3,
    "8",
    3,
    "105",
    2,
    "107",
    2,
    "2",
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
                2,
                "27",
                2,
                "62",
                2,
                "112",
                1,
                "117",
                1,
                "25",
                1,
                "31",
                1,
                "33",
                1,
                "44",
                1
            ],
            "owner_org_title": [
                "National Research Council Canada | Conseil national de recherches
Canada",
                42,
                "Canadian Nuclear Safety Commission | Commission canadienne de
sûreté nucléaire",
                5,
                "National Defence | Défense nationale",
                4,
                "Innovation, Science and Economic Development Canada | Innovation,
Sciences et Développement économique Canada",
                3,
                "Library and Archives Canada | Bibliothèque et Archives Canada",
                3,
                "Agriculture and Agri-Food Canada | Agriculture et Agroalimentaire
Canada",
                2,
                "Communications Security Establishment Canada | Centre de la
sécurité des télécommunications Canada",
                2,
                "Health Canada | Santé Canada",
                2,
                "Infrastructure Canada | Infrastructure Canada",
                2,
                "Natural Resources Canada | Ressources naturelles Canada",
                2,
                "Canadian Space Agency | Agence spatiale canadienne",
                1,
                "Department of Finance Canada | Ministère des Finances Canada",
                1,
                "Environment and Climate Change Canada | Environnement et
Changement climatique Canada",
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
                1,
                "Fisheries and Oceans Canada | Pêches et Océans Canada",
                1,
                "Statistics Canada | Statistique Canada",
                1,
                "Treasury Board of Canada Secretariat | Secrétariat du Conseil du
Trésor du Canada",
                1
            ]
        },
        "facet_ranges": {},
        "facet_intervals": {},
        "facet_heatmaps": {}
    },
    "highlighting": {
        "nrc-cnrc,2022,6,A-2022-09": {
            "disposition": [],
            "comments_en": [],
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_en": [
                "Briefing note BN0007793 Update on the <mark>NRC</mark>’s Clinical
Trial Material Facility"
            ],
            "owner_org_title": []
        },
        "nrc-cnrc,2021,7,A-2021-27": {
            "disposition": [],
            "comments_en": [],
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_en": [
                "Briefing Note Ref # BN0007793 Update on the <mark>NRC</mark>’s
Clinical Trial Material Facility"
            ],
            "owner_org_title": []
        },
        "nrc-cnrc,2022,7,A-2022-14": {
            "disposition": [],
            "comments_en": [],
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_en": [
                "Summary of how many researchers use each of the <mark>NRC</mark>’s
networks (Black network, orange network and green network) as of the March 18, 2022
cyber incident reported by the <mark>NRC</mark>. "
            ],
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
            "owner_org_title": []
        },
        "nrc-cnrc,2022,8,A-2022-05": {
            "disposition": [],
            "comments_en": [],
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_en": [
                "All internal notes and communications at <mark>NRC</mark>, as well
as with other federal organizations, regarding the establishment of the effective
date of <mark>NRC</mark>'s mandatory vaccination policy for November 8, 2021."
            ],
            "owner_org_title": []
        },
        "nrc-cnrc,2022,2,A-2021-38": {
            "disposition": [],
            "comments_en": [],
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_en": [
                "<mark>NRC</mark>'s licensing agreement with CanSino for the HK293
cell line (and any related contracts/agreements/extensions)"
            ],
            "owner_org_title": []
        },
        "nrc-cnrc,2021,8,A-2021-42": {
            "disposition": [],
            "comments_en": [],
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_en": [
                "Briefing note Ref # BN21-06-105, Approval of the Evaluation of
<mark>NRC</mark>'s Metrology Research Centre"
            ],
            "owner_org_title": []
        },
        "nrc-cnrc,2021,10,A-2021-01": {
            "disposition": [],
            "comments_en": [],
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_en": [
                "Documents related to the National Research Council of Canada
(<mark>NRC</mark>) decision to transfer ownership of equipment purchased by
<mark>NRC</mark> to Lanaupôle Fibers Inc. "
            ],
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
            "owner_org_title": []
        },
        "nrc-cnrc,2021,12,A-2021-33": {
            "disposition": [],
            "comments_en": [],
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_en": [
                "Briefing note BN0008318 Support to Invoke National Security
Exception for Timely Design and Build of <mark>NRC</mark>'s Clinical Trial Material
Facility"
            ],
            "owner_org_title": []
        },
        "nrc-cnrc,2022,10,A-2022-08": {
            "disposition": [],
            "comments_en": [],
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_en": [
                "Briefing note BN0008318Support to Invoke National Security
Exception for Timely Design and Build of <mark>NRC</mark>'s Clinical Trial Material
Facility"
            ],
            "owner_org_title": []
        },
        "nrc-cnrc,2022,7,A-2022-12": {
            "disposition": [],
            "comments_en": [],
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_en": [
                "A copy of the email sent by <mark>NRC</mark> President Iain
Stewart to @<mark>NRC</mark> email addresses in March regarding the cyber incident
that affected the National Research Council on March 18, 2022. All other subsequent
emails from Iain Stewart/<mark>NRC</mark> managers or directors to <mark>NRC</mark>
staff/researchers/clients providing updates on the above cyber incident since the
original email. "
            ],
            "owner_org_title": []
        }
    }
}
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
{
    "responseHeader": {
        "status": 0,
        "QTime": 16,
        "params": {
            "f.year.facet.sort": "index",
            "f.month.facet.limit": "12",
            "hl": "on",
            "f.owner_org.facet.limit": "200",
            "indent": "false",
            "fl":
"id,month,month_fr,disposition,disposition_fr,comments_fr,umd_number,owner_org,owne
r_org_fr,year,request_number,summary_fr,pages,pages_fr,owner_org_title",
            "f.umd_number.facet.sort": "count",
            "q.op": "AND",
            "f.owner_org.facet.sort": "label",
            "f.disposition.facet.limit": "50",
            "hl.simple.pre": "<mark>",
            "highlight": "true",
            "defType": "edismax",
            "qf": [
                "id",
                "month",
                "month_fr",
                "disposition",
                "disposition_fr",
                "comments_fr",
                "umd_number",
                "owner_org",
                "owner_org_fr",
                "year",
                "request_number",
                "summary_fr",
                "pages",
                "pages_fr",
                "owner_org_title"
            ],
            "sow": "true",
            "f.umd_number.facet.limit": "100",
            "hl.fl": [
                "disposition",
                "comments_fr",
                "umd_number",
                "owner_org",
                "request_number",
                "summary_fr",
                "owner_org_title"
            ],
            "f.disposition.facet.sort": "count",
            "wt": "json",
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
            "hl.highlightMultiTerm": "true",
            "facet.field": [
                "owner_org",
                "disposition",
                "year",
                "month",
                "umd_number",
                "owner_org_title"
            ],
            "start": "0",
            "f.owner_org_title.facet.sort": "count",
            "f.owner_org_title.facet.limit": "100",
            "sort": "score desc",
            "rows": "10",
            "hl.snippets": "10",
            "q": "nrc",
            "hl.simple.post": "</mark>",
            "f.month.facet.sort": "index",
            "f.year.facet.limit": "100",
            "facet.method": "enum",
            "hl.method": "unified",
            "facet.mincount": "1",
            "facet": "true",
            "facet.sort": "index"
        }
    },
    "response": {
        "numFound": 10,
        "start": 0,
        "numFoundExact": true,
        "docs": [
            {
                "year": 2022,
                "month": 6,
                "month_fr": "juin",
                "request_number": "A-2021-78",
                "summary_fr": "Protocole d’entente Killam-<mark>NRC</mark>-Conseil
des arts du Canada : protocole d’entente concernant l’administration des fiducies
Killam signé en mai 2021.",
                "disposition": "DP",
                "disposition_fr": "Communication partielle",
                "pages": 13,
                "pages_fr": "13",
                "comments_fr": "-",
                "umd_number": "80",
                "owner_org": "nrc-cnrc",
                "owner_org_fr": "Conseil national de recherches Canada",
                "owner_org_title": "National Research Council Canada | Conseil
national de recherches Canada",
                "id": "nrc-cnrc,2022,6,A-2021-78"
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
            },
            {
                "year": 2022,
                "month": 7,
                "month_fr": "juillet",
                "request_number": "A-2020-71",
                "summary_fr": "Tous les courriels et information liés à la décision
du Comité exécutif (CE) de la Commission canadienne des codes du bâtiment et de
prévention des incendies (CCCBPI) de supprimer l'exigence obligatoire d'étanchéité
à l'air du Code national du bâtiment et du Code national de l'énergie pour les
bâtiments ; connecté à : Michel.Dumoulin@<mark>nrc</mark>-cnrc.gc.ca,
Anne.Gribbon@<mark>nrc</mark>-cnrc.gc.ca, Andre.Laroche@<mark>nrc</mark>-cnrc.gc,
ruth.talbot@canada.ca, demian.chaiquin@canada .ca, jim.comtois@canada.ca,
ian.meredith@canada.ca, CCCBPI (Douglas Crawford, Robert Rymell), @chba.ca,
@cement.ca, mattamycorp.com, @canadianchemistry.ca, @minto. com, @cisc-icca.ca,
@cwc.ca. ",
                "disposition": "DP",
                "disposition_fr": "Communication partielle",
                "pages": 786,
                "pages_fr": "786",
                "comments_fr": "-",
                "umd_number": "80",
                "owner_org": "nrc-cnrc",
                "owner_org_fr": "Conseil national de recherches Canada",
                "owner_org_title": "National Research Council Canada | Conseil
national de recherches Canada",
                "id": "nrc-cnrc,2022,7,A-2020-71"
            },
            {
                "year": 2020,
                "month": 3,
                "month_fr": "mars",
                "request_number": "A-2019-00140",
                "summary_fr": "Accords internationaux entre les États-Unis et le
Canada concernant le partage des documents qui couvrent la sûreté nucléaire, la
sécurité publique et la fiabilité du réseau nord-américain; ces documents
permettent aux citoyens canadiens d’obtenir des documents de la <mark>NRC</mark>
des États-Unis et aux citoyens américains d’obtenir des documents de la
CCSN.\n\nNouvelle portée le 13 mars 2020 : Copies des protocoles d’entente et
autres accords entre la CCSN et la <mark>NRC</mark> des États-Unis.",
                "disposition": "NE",
                "disposition_fr": "Aucun document n’existe",
                "pages": 22,
                "pages_fr": "22",
                "comments_fr": "-",
                "umd_number": "29",
                "owner_org": "cnsc-ccsn",
                "owner_org_fr": "Commission canadienne de sûreté nucléaire",
                "owner_org_title": "Canadian Nuclear Safety Commission | Commission
canadienne de sûreté nucléaire",
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
            "id": "cnsc-ccsn,2020,3,A-2019-00140"
        },
        {
            "year": 2021,
            "month": 1,
            "month_fr": "janvier",
            "request_number": "A-2020-00034",
            "summary_fr": "AIPRP : Rapport des représentants et contributeurs
canadiens\nabréviation IRS : Système de notification des incidents pour les
installations nucléaires, AIEA\n\nDeux rapports :\n1) Rapport #1: US
<mark>NRC</mark> Report Number ML13063A135 IRS Meeting\n2) Rapport #2: IAEA J8 CS -
43728\nLIMITED DISTRIBUTION 16 APRIL 2012\nHIGHLIGHTS FROM\nTHE INTERNATIONAL
REPORTING SYSTEM FOR OPERATING EXPERIENCE (IRS) FOR EVENTS IN 2010 –
\n2011\nCoordonnateur du pays pour le système IRS, dans au moins un cas : M. ",
            "disposition": "EX",
            "disposition_fr": "Exception totale",
            "pages": 0,
            "pages_fr": "0",
            "comments_fr": "-",
            "umd_number": "29",
            "owner_org": "cnsc-ccsn",
            "owner_org_fr": "Commission canadienne de sûreté nucléaire",
            "owner_org_title": "Canadian Nuclear Safety Commission | Commission
canadienne de sûreté nucléaire",
            "id": "cnsc-ccsn,2021,1,A-2020-00034"
        },
        {
            "year": 2017,
            "month": 12,
            "month_fr": "decembre",
            "request_number": "A-2017-00133",
            "summary_fr": "Cette demande exige d’effectuer « une analyse
comparative entre la CCSN et la <mark>NRC</mark> des États-Unis concernant les
inspections des centrales nucléaires : le temps consacré à la formation des
inspecteurs; le temps des inspections à chaque centrale; le nombre d’inspections et
le nombre d’inspecteurs ».",
            "disposition": "NE",
            "disposition_fr": "Aucun document n’existe",
            "pages": 0,
            "pages_fr": "0",
            "comments_fr": "-",
            "umd_number": "29",
            "owner_org": "cnsc-ccsn",
            "owner_org_fr": "Commission canadienne de sûreté nucléaire",
            "owner_org_title": "Canadian Nuclear Safety Commission | Commission
canadienne de sûreté nucléaire",
            "id": "cnsc-ccsn,2017,12,A-2017-00133"
        },
        {
            "year": 2013,
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
                "month": 12,
                "month_fr": "decembre",
                "request_number": "A-2013-00428",
                "summary_fr": "Reference   RG29, Accession 1982-83/109, Box 10,
File Number 110-83/109, Part 6, File Title: RPB Committees, Councils and
Conferences - <mark>NRC</mark> Scientific Criteria - Physical Energy Phenomena File
part of: Records from the quarantine and public health registry system, Department
of Agriculture, MIKAN 193532",
                "disposition": "DA",
                "disposition_fr": "Communication totale",
                "pages": 366,
                "pages_fr": "366",
                "comments_fr": "-",
                "umd_number": "8",
                "owner_org": "lac-bac",
                "owner_org_fr": "Bibliothèque et Archives Canada",
                "owner_org_title": "Library and Archives Canada | Bibliothèque et
Archives Canada",
                "id": "lac-bac,2013,12,A-2013-00428"
            },
            {
                "year": 2022,
                "month": 7,
                "month_fr": "juillet",
                "request_number": "A-2022-12",
                "summary_fr": "Une copie du courriel envoyé par le président du
CNRC, Iain Stewart, aux adresses de courriel @<mark>NRC</mark> en mars concernant
l'incident cybernétique qui a touché le Conseil national de recherches du Canada le
18 mars 2022. ",
                "disposition": "DP",
                "disposition_fr": "Communication partielle",
                "pages": 42,
                "pages_fr": "42",
                "comments_fr": "-",
                "umd_number": "80",
                "owner_org": "nrc-cnrc",
                "owner_org_fr": "Conseil national de recherches Canada",
                "owner_org_title": "National Research Council Canada | Conseil
national de recherches Canada",
                "id": "nrc-cnrc,2022,7,A-2022-12"
            },
            {
                "year": 2017,
                "month": 8,
                "month_fr": "août",
                "request_number": "A-2016-00286",
                "summary_fr": "VEUILLEZ FOURNIR DES RENSEIGNEMENTS SUR LES
COMPARAISONS FAITES PAR LE PERSONNEL DE LA CCSN ENTRE LES EXIGENCES CANADIENNES ET
LES EXIGENCES DU GUIDE DE RÉGLEMENTATION 1.9 DE LA <mark>NRC</mark> DES ÉTATS-UNIS,
APPLICATION AND TESTING OF SAFETY-RELATED DIESEL GENERATORS IN NUCLEAR POWER
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

PLANTS.",
                "disposition": "DP",
                "disposition_fr": "Communication partielle",
                "pages": 2855,
                "pages_fr": "2 855",
                "comments_fr": "-",
                "umd_number": "29",
                "owner_org": "cnsc-ccsn",
                "owner_org_fr": "Commission canadienne de sûreté nucléaire",
                "owner_org_title": "Canadian Nuclear Safety Commission | Commission canadienne de sûreté nucléaire",
                "id": "cnsc-ccsn,2017,8,A-2016-00286"
            },
            {
                "year": 2020,
                "month": 1,
                "month_fr": "janvier",
                "request_number": "A-2019-00062",
                "summary_fr": "Cette demande vise des documents relatifs à la sûreté, aux examens techniques ou au processus d'homologation des colis de transport destinés au combustible usé, aux cœurs ou aux composants de réacteur nucléaire (aussi appelé combustible épuisé or combustible irradié ou déchets radioactifs de haute activité), y compris des morceaux et fragments, ainsi qu'aux communications liées à ces examens réalisés par la CCSN entre la CCSN et la SGDN ou ses représentants, entre la CCSN et les Laboratoires Nucléaires Canadiens (LNC) ou leurs représentants, et entre la CCSN et d'autres organismes de réglementation, comme la <mark>NRC</mark> des États-Unis. Les documents demandés visent deux processus d'homologation : 1. ",
                "disposition": "DP",
                "disposition_fr": "Communication partielle",
                "pages": 620,
                "pages_fr": "620",
                "comments_fr": "-",
                "umd_number": "29",
                "owner_org": "cnsc-ccsn",
                "owner_org_fr": "Commission canadienne de sûreté nucléaire",
                "owner_org_title": "Canadian Nuclear Safety Commission | Commission canadienne de sûreté nucléaire",
                "id": "cnsc-ccsn,2020,1,A-2019-00062"
            },
            {
                "year": 2020,
                "month": 1,
                "month_fr": "janvier",
                "request_number": "A-2019-00061",
                "summary_fr": "Cette demande vise des documents relatifs à la sûreté, aux examens techniques ou au processus d'homologation des colis de transport destinés au combustible usé ainsi qu'aux communications liées à ces examens réalisés par la CCSN entre la CCSN et la SGDN ainsi qu'entre la CCSN et d'autres organismes de réglementation, comme la Nuclear Regulatory Commission

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

(<mark>NRC</mark>) des États-Unis. Les documents demandés visent deux processus d’homologation : 1. ",
                "disposition": "DP",
                "disposition_fr": "Communication partielle",
                "pages": 125,
                "pages_fr": "125",
                "comments_fr": "-",
                "umd_number": "29",
                "owner_org": "cnsc-ccsn",
                "owner_org_fr": "Commission canadienne de sûreté nucléaire",
                "owner_org_title": "Canadian Nuclear Safety Commission | Commission canadienne de sûreté nucléaire",
                "id": "cnsc-ccsn,2020,1,A-2019-00061"
            }
        ]
    },
    "facet_counts": {
        "facet_queries": {},
        "facet_fields": {
            "owner_org": [
                "cnsc-ccsn",
                6,
                "lac-bac",
                1,
                "nrc-cnrc",
                3
            ],
            "disposition": [
                "DP",
                6,
                "NE",
                2,
                "DA",
                1,
                "EX",
                1
            ],
            "year": [
                "2013",
                1,
                "2017",
                2,
                "2020",
                3,
                "2021",
                1,
                "2022",
                3
            ],
            "month": [

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
                "1",
                3,
                "3",
                1,
                "6",
                1,
                "7",
                2,
                "8",
                1,
                "12",
                2
            ],
            "umd_number": [
                "29",
                6,
                "80",
                3,
                "8",
                1
            ],
            "owner_org_title": [
                "Canadian Nuclear Safety Commission | Commission canadienne de
sûreté nucléaire",
                6,
                "National Research Council Canada | Conseil national de recherches
Canada",
                3,
                "Library and Archives Canada | Bibliothèque et Archives Canada",
                1
            ]
        },
        "facet_ranges": {},
        "facet_intervals": {},
        "facet_heatmaps": {}
    },
    "highlighting": {
        "nrc-cnrc,2022,6,A-2021-78": {
            "disposition": [],
            "comments_fr": [],
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_fr": [
                "Protocole d’entente Killam-<mark>NRC</mark>-Conseil des arts du
Canada : protocole d’entente concernant l’administration des fiducies Killam signé
en mai 2021."
            ],
            "owner_org_title": []
        },
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
        "nrc-cnrc,2022,7,A-2020-71": {
            "disposition": [],
            "comments_fr": [],
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_fr": [
                "Tous les courriels et information liés à la décision du Comité
exécutif (CE) de la Commission canadienne des codes du bâtiment et de prévention
des incendies (CCCBPI) de supprimer l'exigence obligatoire d'étanchéité à l'air du
Code national du bâtiment et du Code national de l'énergie pour les bâtiments ;
connecté à : Michel.Dumoulin@<mark>nrc</mark>-cnrc.gc.ca,
Anne.Gribbon@<mark>nrc</mark>-cnrc.gc.ca, Andre.Laroche@<mark>nrc</mark>-cnrc.gc,
ruth.talbot@canada.ca, demian.chaiquin@canada .ca, jim.comtois@canada.ca,
ian.meredith@canada.ca, CCCBPI (Douglas Crawford, Robert Rymell), @chba.ca,
@cement.ca, mattamycorp.com, @canadianchemistry.ca, @minto. com, @cisc-icca.ca,
@cwc.ca. "
            ],
            "owner_org_title": []
        },
        "cnsc-ccsn,2020,3,A-2019-00140": {
            "disposition": [],
            "comments_fr": [],
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_fr": [
                "Accords internationaux entre les États-Unis et le Canada
concernant le partage des documents qui couvrent la sûreté nucléaire, la sécurité
publique et la fiabilité du réseau nord-américain; ces documents permettent aux
citoyens canadiens d’obtenir des documents de la <mark>NRC</mark> des États-Unis et
aux citoyens américains d’obtenir des documents de la CCSN.\n\nNouvelle portée le
13 mars 2020 : Copies des protocoles d’entente et autres accords entre la CCSN et
la <mark>NRC</mark> des États-Unis."
            ],
            "owner_org_title": []
        },
        "cnsc-ccsn,2021,1,A-2020-00034": {
            "disposition": [],
            "comments_fr": [],
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_fr": [
                "AIPRP : Rapport des représentants et contributeurs
canadiens\nabréviation IRS : Système de notification des incidents pour les
installations nucléaires, AIEA\n\nDeux rapports :\n1) Rapport #1: US
<mark>NRC</mark> Report Number ML13063A135 IRS Meeting\n2) Rapport #2: IAEA J8 CS -
43728\nLIMITED DISTRIBUTION 16 APRIL 2012\nHIGHLIGHTS FROM\nTHE INTERNATIONAL
REPORTING SYSTEM FOR OPERATING EXPERIENCE (IRS) FOR EVENTS IN 2010 –
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

\n2011\nCoordonnateur du pays pour le système IRS, dans au moins un cas : M. ",
                "Poulet\nDossiers de la base de données pour le rapport IRS 8096;\nDossiers de la base de données pour le rapport #1: US <mark>NRC</mark> Report Number ML13063A135 IRS Meeting\nDossiers de la base de données pour le rapport #2: IAEA J8 CS – 43728\nCourriel(s) ou toute autre forme de communication entre le coordonnateur de pays de la CCSN pour le système IRS et l'Office of Nuclear Reactor Regulation de la U.S.\nNuclear Regulatory Commission (US <mark>NRC</mark>) ou des agents comme Mme\nRebecca Sigmon, courriel : Rebecca.Sigmon@nrc.gov.\nÉchange de courriels ou toute autre forme de communication entre le coordonnateur de pays de la CCSN et des membres du personnel de la CCSN concernant le rapport RS 8096, le rapport #1:US <mark>NRC</mark> Report Number ML13063A135 IRS Meeting et le rapport #2: IAEA J8 CS – 43728.\nNuméros edoc des rapports suivants :\nIRS 8096,\nRapport #1:US <mark>NRC</mark> Report Number ML13063A135 IRS Meeting and\nRapport #2: IAEA J8 CS – 43728.\n",
                "Comme mentionné précédemment, les rapports sont les suivants :\nRapport #1: US <mark>NRC</mark> Report Number ML13063A135 IRS Meeting\n\nRapport #2: IAEA J8 CS - 43728\nLIMITED DISTRIBUTION 16 APRIL 2012\nHIGHLIGHTS FROM\nTHE INTERNATIONAL REPORTING SYSTEM FOR OPERATING EXPERIENCE (IRS) FOR EVENTS IN 2010 –\n2011\nCourriels\nÉchanges de courriels ou toute autre forme de communication avec M. François Rinfret ou le personnel en charge des rapports au sujet des deux rapports, comme suit :\nRapport #1: US <mark>NRC</mark> Report Number ML13063A135 IRS Meeting\nRapport #2: IAEA J8 CS - 43728\nLIMITED DISTRIBUTION 16 APRIL 2012\nHIGHLIGHTS FROM\nTHE INTERNATIONAL REPORTING SYSTEM FOR OPERATING EXPERIENCE (IRS) FOR EVENTS IN 2010 –\n2011.\n",
                "Heppell-Masys ou le personnel chargé des rapports au sujet des deux rapports suivants :\nRapport #1: US <mark>NRC</mark> Report Number ML13063A135 IRS Meeting\nRapport #2: IAEA J8 CS - 43728\nLIMITED DISTRIBUTION 16 APRIL 2012\nHIGHLIGHTS FROM\nTHE INTERNATIONAL REPORTING SYSTEM FOR OPERATING EXPERIENCE (IRS) FOR EVENTS IN 2010 –\n2011."
            ],
            "owner_org_title": []
        },
        "cnsc-ccsn,2017,12,A-2017-00133": {
            "disposition": [],
            "comments_fr": [],
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_fr": [
                "Cette demande exige d'effectuer « une analyse comparative entre la CCSN et la <mark>NRC</mark> des États-Unis concernant les inspections des centrales nucléaires : le temps consacré à la formation des inspecteurs; le temps des inspections à chaque centrale; le nombre d'inspections et le nombre d'inspecteurs »."
            ],
            "owner_org_title": []
        },
        "lac-bac,2013,12,A-2013-00428": {
            "disposition": [],
            "comments_fr": [],

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*


```
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_fr": [
                "Reference   RG29, Accession 1982-83/109, Box 10, File Number
110-83/109, Part 6, File Title: RPB Committees, Councils and Conferences -
<mark>NRC</mark> Scientific Criteria - Physical Energy Phenomena File part of:
Records from the quarantine and public health registry system, Department of
Agriculture, MIKAN 193532"
            ],
            "owner_org_title": []
        },
        "nrc-cnrc,2022,7,A-2022-12": {
            "disposition": [],
            "comments_fr": [],
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_fr": [
                "Une copie du courriel envoyé par le président du CNRC, Iain
Stewart, aux adresses de courriel @<mark>NRC</mark> en mars concernant l'incident
cybernétique qui a touché le Conseil national de recherches du Canada le 18 mars
2022. "
            ],
            "owner_org_title": []
        },
        "cnsc-ccsn,2017,8,A-2016-00286": {
            "disposition": [],
            "comments_fr": [],
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_fr": [
                "VEUILLEZ FOURNIR DES RENSEIGNEMENTS SUR LES COMPARAISONS FAITES
PAR LE PERSONNEL DE LA CCSN ENTRE LES EXIGENCES CANADIENNES ET LES EXIGENCES DU
GUIDE DE RÉGLEMENTATION 1.9 DE LA <mark>NRC</mark> DES ÉTATS-UNIS, APPLICATION AND
TESTING OF SAFETY-RELATED DIESEL GENERATORS IN NUCLEAR POWER PLANTS."
            ],
            "owner_org_title": []
        },
        "cnsc-ccsn,2020,1,A-2019-00062": {
            "disposition": [],
            "comments_fr": [],
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_fr": [
                "Cette demande vise des documents relatifs à la sûreté, aux examens
techniques ou au processus d’homologation des colis de transport destinés au
combustible usé, aux cœurs ou aux composants de réacteur nucléaire (aussi appelé
```

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

combustible épuisé or combustible irradié ou déchets radioactifs de haute activité), y compris des morceaux et fragments, ainsi qu’aux communications liées à ces examens réalisés par la CCSN entre la CCSN et la SGDN ou ses représentants, entre la CCSN et les Laboratoires Nucléaires Canadiens (LNC) ou leurs représentants, et entre la CCSN et d’autres organismes de réglementation, comme la <mark>NRC</mark> des États-Unis. Les documents demandés visent deux processus d’homologation : 1. "
            ],
            "owner_org_title": []
        },
        "cnsc-ccsn,2020,1,A-2019-00061": {
            "disposition": [],
            "comments_fr": [],
            "umd_number": [],
            "owner_org": [],
            "request_number": [],
            "summary_fr": [
                "Cette demande vise des documents relatifs à la sûreté, aux examens techniques ou au processus d’homologation des colis de transport destinés au combustible usé ainsi qu’aux communications liées à ces examens réalisés par la CCSN entre la CCSN et la SGDN ainsi qu’entre la CCSN et d’autres organismes de réglementation, comme la Nuclear Regulatory Commission (<mark>NRC</mark>) des États-Unis. Les documents demandés visent deux processus d’homologation : 1. "
            ],
            "owner_org_title": []
        }
    }
}

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Lauzon, Dominique (he/him, il)

From: Trew, Debbie
Sent: September 27, 2023 12:51 PM
To: Fleischer, Leonard
Cc: Mockett, Ryan; Herbert, Wesley; Vadnais, Philippe
Subject: Change to Search

UNCLASSIFIED / NON CLASSIFIÉ

Hello
Please find attached a new screen designed by Wes to bypass our current Search. If Lixar turns off our search next week we will have to go to the OG Portal to do a search against summaries. I am currently getting the one screen translated. If you could share with the developers and let me know if they have any questions or concerns that would be great. Dom is asking that we be ready to implement the change by next Tuesday if we need to. I am currently waiting to hear back from Lixar/BDO to see if they will turn us off.

https://preview.uxpin.com/1a528dc56d1b452725bac786516a57f4d6d237cf#/p/164403436/untitled/page-1

Regards
Debbie Trew
Senior Project Officer
Open Government, Office of the Chief Information Officer
Treasury Board of Canada Secretariat / Government of Canada
Debbie.Trew@tbs-sct.gc.ca / 613-219-9626

Responsable principal de projet
Division du gouvernement overt, Bureau du dirigeant principal de l'information
Secrétariat du Conseil du Trésor du Canada / Gouvernement du Canada
Debbie.Trew@tbs-sct.gc.ca / 613-219-9626

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Lauzon, Dominique (he/him, il)

From: Annett, Megan
Sent: September 27, 2023 2:37 PM
To: Trew, Debbie; Sajjad, Rabia; Herbert, Wesley
Cc: Thompson, Ross; Noble, Paul
Subject: RE: Code uploaded to Global Portal

Thanks Debbie and Rabia. A reminder that we have two long weekends in a row: Oct 2nd and 9th.

From: Trew, Debbie
Sent: Wednesday, September 27, 2023 2:33 PM
To: Sajjad, Rabia ; Herbert, Wesley
Cc: Annett, Megan ; Thompson, Ross ; Noble, Paul
Subject: RE: Code uploaded to Global Portal

UNCLASSIFIED / NON CLASSIFIÉ

Hi Rabia
Thanks for all the information. You were very clear on what work will need to be done. Thanks for help on this.

Update on Search - there is a possibility that we will need to do the Search change before Ross gets back, possibly next week. We are currently working on a new Search screen for the AORS. We wont know for sure until we hear back from GCStrategies. I will let you know as soon as I hear something.

Regards

Debbie Trew
Senior Project Officer
Open Government, Office of the Chief Information Officer
Treasury Board of Canada Secretariat / Government of Canada
Debbie.Trew@tbs-sct.gc.ca / 613-219-9626

Responsable principal de projet
Division du gouvernement ouvert, Bureau du dirigeant principal de l'information
Secrétariat du Conseil du Trésor du Canada / Gouvernement du Canada
Debbie.Trew@tbs-sct.gc.ca / 613-219-9626

From: Sajjad, Rabia <Rabia.Sajjad@tbs-sct.gc.ca>
Sent: Wednesday, September 27, 2023 2:24 PM
To: Trew, Debbie <Debbie.Trew@tbs-sct.gc.ca>; Herbert, Wesley <Wesley.Herbert@tbs-sct.gc.ca>
Cc: Annett, Megan <Megan.Annett@tbs-sct.gc.ca>; Thompson, Ross <Ross.Thompson@tbs-sct.gc.ca>; Noble, Paul <Paul.Noble@tbs-sct.gc.ca>
Subject: RE: Code uploaded to Global Portal

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

s.19(1)

He Debbie,

Might be worthwhile explaining a bit more – migrating the recommendation system as Azure Cloud Apps is no longer an option because IMTD is not available to do that.

Ross' approach now is to use the codebase for recommendation system to re-create the python environment and try to make it work on a similar infrastructure as ours. In addition to that, he will try to simplify the code, if possible.

Hope I haven't created confusion with the term "migration".

Thanks,
Rabia

From: Sajjad, Rabia
Sent: Wednesday, September 27, 2023 1:59 PM
To: Trew, Debbie <Debbie.Trew@tbs-sct.gc.ca>; Herbert, Wesley <Wesley.Herbert@tbs-sct.gc.ca>
Cc: Annett, Megan <Megan.Annett@tbs-sct.gc.ca>; Thompson, Ross <Ross.Thompson@tbs-sct.gc.ca>; Noble, Paul <Paul.Noble@tbs-sct.gc.ca>
Subject: RE: Code uploaded to Global Portal

Hi Debbie,

Last week on Friday, Ross was able to download the code for recommendation system from BDO. He did not find any detailed design documentation but came across some documentation to give him an idea of the important components of the codebase. His impression of the code was that it is over-engineered.

Ross is going to [redacted] to resume his work on this. My understanding is that he will prioritize the ATI search and once that is in good shape, he will then move on to migrating the recommendation system.

I have copied Paul on the email so that he can manage the work for upcoming sprints.

Thanks,
Rabia

From: Thompson, Ross <Ross.Thompson@tbs-sct.gc.ca>
Sent: Friday, September 22, 2023 9:14 AM
To: Dostaler, Paul (BDO Lixar) <pdostaler@bdo.ca>; Sajjad, Rabia <Rabia.Sajjad@tbs-sct.gc.ca>
Cc: Trew, Debbie <Debbie.Trew@tbs-sct.gc.ca>; Herbert, Wesley <Wesley.Herbert@tbs-sct.gc.ca>
Subject: RE: Code uploaded to Global Portal

UNCLASSIFIED / NON CLASSIFIÉ

Hi Paul,

Thank you for your help and finding and sharing this with us. Lots of reading for today!

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Best regards,
Ross

IT Team Lead, Office of the Chief Information Officer
Treasury Board of Canada Secretariat / Government of Canada
Ross.Thompson@tbs-sct.gc.ca / Tel: 613.371.4187 / TTY: 613-369-9371

Chef d'équipe TI, Bureau du dirigeant principal de l'information
Secrétariat du Conseil du Trésor du Canada / Gouvernement du Canada
Ross.Thompson@tbs-sct.gc.ca / Tél : 613.371.4187 / ATS : 613-369-9371

From: Dostaler, Paul (BDO Lixar) <pdostaler@bdo.ca>
Sent: Friday, September 22, 2023 8:08 AM
To: Thompson, Ross <Ross.Thompson@tbs-sct.gc.ca>; Sajjad, Rabia <Rabia.Sajjad@tbs-sct.gc.ca>
Cc: Trew, Debbie <Debbie.Trew@tbs-sct.gc.ca>
Subject: Code uploaded to Global Portal

EXTERNAL EMAIL - USE CAUTION / COURRIEL EXTERNE - FAITES PREUVE DE PRUDENCE

Hi Rabia and Ross,

I can see that our team uploaded the code to the portal last night. In the code folder, there is a readme.txt file that explains how to import it to git. We had to compress the files because the portal didn't support the raw file format. I assume that you do have access to a git instance where you can import?

As for the documentation, what I found was not as deep as I had hoped in terms of detailed design. I have not looked at the code repos, but it's possible that some documentation exists there as well. Let me know how you make out, and we can reconnect next week.

Feel free to send any questions my way in the mean time.

Thanks,
Paul

Paul Dostaler, P.Eng
Partner - Industry Lead - Manufacturing
BDO Lixar
Phone: 613-292-8454
pdostaler@bdo.ca

BDO LIXAR

www.bdo.ca
Before you print think about the environment

BDO is a proud sponsor of Golf Canada
BDO est fier de commanditer Golf Canada

BDO is proud to be among Canada's Top 100 Employers
BDO est fier d'être l'un des 100 meilleurs employeurs du Canada

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

The information contained in this communication is confidential and intended only for the use of those to whom it is addressed. If you have received this communication in error, please notify me by telephone (collect if necessary) and delete or destroy any copies of it. Thank you.

To update your commercial electronic message preferences or unsubscribe from promotional messages, please go to our online subscription centre https://go.bdo.ca/subscription-center.html

BDO Canada LLP (and its affiliates), a Canadian limited liability partnership, is a member of BDO International Limited, a UK company limited by guarantee, and forms part of the international BDO network of independent member firms.

BDO is the brand name for the BDO network and for each of the BDO Member Firms.

Le contenu de ce courriel est confidentiel et à l'intention du (des) destinataire(s) seulement. Si vous recevez cette transmission par erreur, veuillez m'aviser immédiatement par téléphone en utilisant le numéro mentionné ci-haut (à frais virés si nécessaire). Veuillez effacer ou détruire toutes copies de ce courriel reçues. Merci de votre collaboration.

Pour mettre à jour vos préférences en matière de messages électroniques ou si vous ne souhaitez recevoir des messages électroniques commerciaux, veuillez-vous rendre au centre d'inscription en ligne.
https://go.bdo.ca/subscription-center.html

BDO Canada s.r.l./S.E.N.C.R.L. (et ses filiales), une société canadienne à responsabilité limitée/société en nom collectif à responsabilité limitée, est membre de BDO International Limited, société de droit anglais, et fait partie du réseau international de sociétés membres indépendantes BDO.

BDO est la marque utilisée pour désigner le réseau BDO et chacune de ses sociétés membres.

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Project Update – October 3, 2023

QA –

IMTD QA resource still working on test plan. As soon as QA becomes available the QA resource will begin testing.

Discussed this morning during our meeting that we need to know when QA issues have been resolved and the system will be reading for testing. They are currently still working on resolving problems with SIC. Shawn Turcotte is now involved and is working to resolve the issue.

We also noted this morning that we would not be doing UAT at the same time as QA. It is critical that they work through all issues before we start our UAT. Ryan is aware of the concerns and will work with IMTD to get some firm timelines.

Backlog grooming –

The team reviewed all items currently sitting in back log and have brought forward at least 10 items for the developers to work on in case they run out of work.

Lixar/BDO –

[redacted] I have copied Dom on the last email sent to Kristian and have asked him to get back to Dom with there response. Keeping my fingers crossed.

[redacted] we have designed a new Search screen which has been given to IMTD.

- The new screen has been completed by IMTD in English and French.
- QA and UAT have been completed on the new screen
- IMTD will be ready to roll out the new screen as soon as well tell them to. It will only take them about 5 minutes to make the change.
- IMTD have also gone to CAB to get permission to roll out the change when its needed. I have asked them to keep the ticket open with CAB until the end of the month.
- Wes and Ryan are aware of all the details for rolling out the new Search screen.

[redacted]

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

CBSA Timelines -

I left this on my report from the last time. This work is still ongoing. Please see timelines below -

Nov 15 – CBSA is removed from the IRCC site (last possible day to get them live)

Nov 9 – fallback date for CBSA to go live in production

Nov 8 – CBSA go live date in prod

Nov 3 – deadline for ADS to sign off on UAT

Oct 25 – CBSA custom questions in UAT environment

Oct 13 – deadline for ADS to provide finalized custom questions with translation to IMTD

Oct 6 – deadline for CBSA to provide finalized custom Qs with translation

Sep 22 – deadline for CBSA to respond to proposed changes to custom Qs by ADS

IMTD are currently waiting for the custom flow text for CBSA. They are also aware of the fact that the new screen must be completed and approved by all before November 15th.

Misc

If you have any project questions, outside of these notes, please reach out to Wes or Ryan. I have tried to make sure that I have pulled at least one of them into any open items I am working on.

The big items that will need to be addressed over the next few weeks –

- UAT of the next release.
- Continue to monitor our backlog in JIRA.
- Finalizing the CBSA custom questions and translating.
- Redesign of the Search screen and search functionality – this work has been started – Phil and Ross working together.
- Redesign of our Recommendation engine and Expedia experience (code from Lixar/BDO has been sent).
- API work still need to be completed.

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Lauzon, Dominique (he/him, il)

Subject:	Jumpstart for AORS Search functionality
Location:	Microsoft Teams Meeting
Start:	Thu 2023-10-19 2:00 PM
End:	Thu 2023-10-19 2:30 PM
Show Time As:	Tentative
Recurrence:	(none)
Meeting Status:	Not yet responded
Organizer:	Herbert, Wesley
Required Attendees:	Thompson, Ross; Vadnais, Philippe; Tremblay, Dominique
Optional Attendees:	Mockett, Ryan

Meeting to jumpstart integration of OG Search into the AORS to replace the existing Lixar search.

Please feel free to forward to any team members that should be in attendance.

Microsoft Teams meeting

Join on your computer, mobile app or room device
Click here to join the meeting

Meeting ID: 224 276 997 227
Passcode: rNvSVR
Download Teams | Join on the web

Or call in (audio only)
+1 343-803-0870,,275670327# Canada, Ottawa-Hull
Phone Conference ID: 275 670 327#
Find a local number | Reset PIN

Learn More | Meeting options

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

s.20(1)(b)

s.20(1)(c)

Lauzon, Dominique (he/him, il)

From: Violette-Fehr, Nancy
Sent: October 26, 2023 12:12 PM
To: Burt, Stephen (he/him, il)
Cc: Langevin, Brigitte; Lauzon, Dominique (he/him, il); Guindon, Marie-France
Subject: ATIP Online Reduced Functionality

Categories: Division / EDO

UNCLASSIFIED / NON CLASSIFIÉ

Hi Stephen,

I am writing to inform you of reduced ATIP Online functionality over the next 2 months. This information will be in the Weekly Operations Report for next week. SCMA, ATI Policy and Privacy Policy have been informed given the visibility of anything related to ATIP.

Scope

- On September 26, 2023, it became necessary to remove the ATI Summaries search tools and AI-enabled recommendation engines from ATIP Online.

Impact

- There will be no impact on the ability of requesters to submit ATI or personal information requests.
- There will be no impact to institutions.
- Requesters will temporarily lose access to optional integrated tools to help guide requesters to the institution(s) most likely to hold sought after records. However, requesters will be given the option to be redirected to the existing ATI Summaries search tool on the Open Government Portal.

Action Plan for Resolution

- October 26: The integrated search and recommendation for ATI requests was replaced with an external link to the ATI Summaries search tool on the Open Government Portal. An emergency patch was required to allow for the continued development of functionality for CBSA's onboarding on November 8.
- November 8: The integrated ATI Summaries search will be restored using a new solution developed internally by the Open Government Portal Team. This will be implemented during a planned update to enable the onboarding of CBSA to ATIP Online.
- Early December: A new more efficient, less costly recommendation engine being developed by the Open Government Portal team will be implemented. As special update will be scheduled as soon as the new tools are developed and fully tested.

Cause

- The impact tools were developed by a third-party (Lixar, now BDO-Lixar) in 2019 and have been hosted by them as a sub-contractor through a contract with GC Strategies.
- The GC Strategies contract (24062-19-488) expired on September 30, 2023.
- [redacted]
- Due vacancies and the inability to staff these as well as unexpected absences, development of replacement tools could not be completed before the end of the extended service.

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Note – While the sequencing of these events is unfortunate and suboptimal, once resolved this solution will reduce our dependency on professional service contracts in exchange better integrated use of internal resources through cross-product collaboration.

Dominique and I are available should have any questions.

Nancy

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Lauzon, Dominique (he/him, il)

From: Lauzon, Dominique (he/him, il)
Sent: October 31, 2023 4:05 PM
To: Burt, Stephen (he/him, il)
Cc: Violette-Fehr, Nancy; Taillefer, Charles
Subject: RE: INFO: Technical issue with Portal Publications on the Open Government Portal

Stephen,

WRT the reduced search functionality within the ATIP Online Request Service (AORS), the replacement ATI summaries search feature has been developed and is currently being tested. We are optimistic that we will be able to roll it out as part of the update to onboard CBSA next week. The result of this is that users will no longer have to leave the application to search the ATI summaries.

The replacement recommendation engine (the tool that identifies the institutions most likely to hold the sought after records) is still being designed. Primary development is expected to be completed by the end of November. Once fully tested, we will push it into production as an emergency update.

Regards,

Dom

From: Burt, Stephen (he/him, il) <Stephen.Burt@tbs-sct.gc.ca>
Sent: Tuesday, October 31, 2023 3:59 PM
To: Lauzon, Dominique (he/him, il) <Dominique.Lauzon@tbs-sct.gc.ca>
Cc: Violette-Fehr, Nancy <Nancy.Violette-Fehr@tbs-sct.gc.ca>; Taillefer, Charles <Charles.Taillefer@tbs-sct.gc.ca>
Subject: RE: INFO: Technical issue with Portal Publications on the Open Government Portal

UNCLASSIFIED / NON CLASSIFIÉ

Thanks Dominique, I agree that this doesn't need to be reported up. Let me know if there is anything you need from me as you work to resolve it.

Separately, any updates on the functionality issue with ATIP Online? If no, that's fine. Just let me know as that one develops as well.

Stephen

From: Lauzon, Dominique (he/him, il) <Dominique.Lauzon@tbs-sct.gc.ca>
Sent: Tuesday, October 31, 2023 2:47 PM
To: Burt, Stephen (he/him, il) <Stephen.Burt@tbs-sct.gc.ca>
Cc: Violette-Fehr, Nancy <Nancy.Violette-Fehr@tbs-sct.gc.ca>; Taillefer, Charles <Charles.Taillefer@tbs-sct.gc.ca>
Subject: INFO: Technical issue with Portal Publications on the Open Government Portal

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Good afternoon Stephen,

Last week, the OG Portal team started receiving reports of Proactive Publication data not getting published. The issue is more complex than initially suspected. However, the situation has been mostly resolved.

ISSUE
Originally, the issue was thought to be related to data uploaded to the Open Government Registry by institutions not getting pushed the data repository of the public Open Government Portal. We have since determined that some data entries include errors or offending strings that interfere with the search algorithms and prevent the proper display of search results.
Is important to note that **no data has been lost**. All data are included in the downloadable data set. The issue is limited to the search results in the user interface.

CONTEXT
Many Proactive Publications, under the ATIA Part 2 (contracts, travel, hospitality…) and policy (E.g. ATI Summaries), must be published on a monthly or quarterly basis. All of these needed to be published by yesterday (within 30 days of the end of the month or quarter).

IMPACT
A very small number of data entries will not appear to public users conducting searches through the user interface. The underlying datasets are complete and are available for download.

CURRENT STATUS
- All policy centres have been informed.
- We have adjusted the search tools for the affect Proactive Publications to limit the search errors to the individual entries that are at cause.
- We are extracting all of the offending data entries to identify the root cause.
- We have added a contextual alert on the affected search screens informing users that "We are experiencing minor technical difficulties with some proactive publications not displaying in the search. We are working to resolve this issue."

The team continues to search for the underlying cause. Once identified, we will work with the data owners to replace the offending data entries.

Given the limited impact and mitigation measures in place, I do not believe that communications products will be necessary.

I am available should you have any questions.

Regards,

Dom

Dominique Lauzon (he/him | il)

a/Senior Director, Office of the Chief Information Officer

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Treasury Board of Canada Secretariat / Government of Canada
Dominique.Lauzon@tbs-sct.gc.ca / Mobile Tel: 613-854-4387

Directeur Principal p.i., Bureau du dirigeant principal de l'information
Secrétariat du Conseil du Trésor du Canada / Gouvernement du Canada
Dominique.Lauzon@tbs-sct.gc.ca / Tél. portable: 613-854-4387

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Lauzon, Dominique (he/him, il)

From: Trew, Debbie
Sent: October 31, 2023 11:40 AM
To: Annett, Megan
Cc: Herbert, Wesley
Subject: Recommendation Engine
Attachments: GC Strategies Inc CT# 24062-17-168 TA for Exploratory Initiatives for Recommendation engine.doc

UNCLASSIFIED / NON CLASSIFIÉ

Hello
I don't have a lot of information on the recommendation engine, but I did find a copy of the TA that explains in some details what Lixar was planning to do. Let me know if this helps at all? If not I will keep looking around to see what I can find.

Debbie Trew
Senior Project Officer
Open Government, Office of the Chief Information Officer
Treasury Board of Canada Secretariat / Government of Canada
Debbie.Trew@tbs-sct.gc.ca / 613-219-9626

Responsable principal de projet
Division du gouvernement overt, Bureau du dirigeant principal de l'information
Secrétariat du Conseil du Trésor du Canada / Gouvernement du Canada
Debbie.Trew@tbs-sct.gc.ca / 613-219-9626

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

TASK AUTHORIZATION	
Contractor: GC Strategies / Lixar	**Contract Number: 24062-17-168**
Commitment Number:	**Financial Coding: 200504**
Task Number: 17168TA-	**Date: May 1, 2018**
TA Request (For completion by Technical Authority)	

TA Request for Exploratory Initiatives for Recommendation Engine

Background/Objective:

The Government committed to create a simple, central website where Canadians can submit access to information and personal information requests to any government institution. The new central website will build on the scope of capabilities currently available in the ATIP Online Pilot, add to them, make them available to over 240 institutions, and enhance the user experience.

Scope of Work:

Three (3) Data Conversion Specialists (I1.0) – Level 3

- Four exploratory initiatives. Each initiative's objectives are to provide incremental improvements to the Departmental Recommendations.
- The Open Data Exploration, Relevance Feedback and System Logs concepts will explore different sources of data and evaluate their impacts on the model's performance.
- The Sufficient Information will assess whether it's possible to measure what a "**good**" search criteria looks like. This will follow a slightly different approach, as such will be detailed separately from the previous 3.

Tasks and Deliverables

Sufficient Information

- The project will seek to measure when an appropriately descriptive input string has been entered and guide the user to improving his terms so that more accurate results can be returned to him. By looking at the results probability distribution, it should provide a hint as to how useful the terms used were in locating something of value. For example, terms yielding 100 results, each with a 1% probability that it is relevant to the user, would essentially translate to "I couldn't find anything of relevance, so here's a bunch of results with equal relevance". Conversely, if the 100 results had 1 with 95% relevance, and the other 99 were 0.05%, we might assume that the 95% relevance one is very relevant.
- Explore options to indicate how performant the entered terms are so that we are confident the recommendations are as meaningful to the user as possible. If the system detects low confidence in the entered terms, the user can be informed so that he may add in additional terms to further define what they're looking for.
- Methodology
- Department Recommendations
- Trying to find a threshold for the sum of the probabilities of the top 5 results.
- Treat this as a binary classification problem where the top 5 is one class, and the rest are our second class

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

- Using the ROC curve and the F-Measure, identify the threshold that would maximize our accuracy based on user input. Call this a confidence threshold, where if the sum of the probabilities of the top 5 is above this value, then be certain that the user has provided sufficient information
- If the sum (based on user input) falls below the threshold, ask the user to provide more information
- Summary Search
- Threshold on the cosine similarity of user input and retrieved documents
- Deliverable
- A document detailing the experiment's respective results, with a recommendation on the way forward which would include identified thresholds.

Relevance Feedback

- This initiative would look to leverage user behaviour into the model in the hopes that it would increase overall accuracy. The idea would be that a successful recommendation would lead to fewer erratic user interactions, whereas poor recommendations may have the user navigating back and forth and/or attempting many variations of similar wording.
- Additionally, inspecting the end to end user flow and evaluating success or failure could further enhance the relevance feedback. For existing summaries, a user flow ending in the initiation of an informal request of an existing ATIP report would be considered positive. For departmental recommendations, the end to end flow of a user navigating to a recommended department, submitting an ATIP request, and having that ATIP request be properly provisioned would also be considered a positive signal.
- The focus for this concept will be to review erratic user behaviour and evaluate whether this leads to improved performance.
- Methodology
- Relevance feedback is the process of search parameter adjustment following the initial search result retrieval. In this process, users' behaviour following the initial retrieval is analyzed to determine the relevance of the results to the query (e.g. clicking the back button, looking at multiple results, etc.). In many cases, this data can be leveraged to better optimize the engine.
- Conduct an exploratory proof of concept around the hypothesis that capturing simple user behaviour on the UI layer can create data for the Relevance Feedback model which can then be leveraged to generate more accurate results.
- Deliverable
- The results will be validated through a series of simulations and will be provided in a summary report. This report will also contain high level instructions on recommended next steps to implement the improvement.

Relevance Feedback - System Logs

- If incorporating relevance feedback shows promising results, it's possible the system

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

logs will be a good place to look for historical traces of user flows demonstrating positive or negative signals. Assuming these logs exist and IMTD can retrieve logs for many years of activity, look to re-create a structured view of individual user workflows from these system logs. These flows could then be used in the relevance feedback model to further improve the overall system performance and accuracy.

- This exploratory proof of concept will test the hypothesis that system logs (assuming they are not Protected B) could be used to reinforce positive or negative signals.
- Methodology
- Assuming system logs will be coming from multiple different sources, it won't be possible to perfectly match user flows between log systems. A fuzzy matching strategy will be developed to gather up likely user flows together, using whatever data is available from the logs. For example, timestamps and source IP's may be sufficient to create a view of the user's navigation.
- Once the flows have been modeled, match them to the datasets the relevance feedback model expects and test out its impact on the overall system performance.
- Deliverable
- A document outlining the experiments and its results. If successful, detailed descriptions on the fuzzy matching strategy as well as high level implementation plan to bring this strategy to production.

Open Data Exploration

- There is a wealth of data on the OpenGov site. Skim through the datasets looking for any potential data that could be used to increase the overall accuracy and robustness of the predictive models.
- In parallel, a visualization team will identify a small handful of interesting datasets and explore them through data visualization. Ultimately, looking for either an informative data story or some potential for future functionality or concepts that could be added to the TBS AI Platform.
- Deliverable
- Depending on what analysts find, either deliver a report recommending the inclusion of certain datasets into the models for improved accuracy or deliver a data visualization highlighting an informative data story which could lead to future platform improvements.

2. PERIOD OF SERVICES	**From:**	**Date of award**	**To:**	**March 31st, 2018**
3. Work Location	Work to be done at suppliers office, with occasional TBS on-site meetings			
4. Travel Requirements	**NO**			
5. Other Conditions /Restraints	**NO**			
6. Task Proposal (insert rows as required) Check (☐):	**Estimated Cost** ☐	**Fixed Price X**	$_____	
7. LEVEL OF SECURITY CLEARANCE REQUIRED FOR THE CONTRACTOR'S PERSONNEL				
Reliability Status				
8. BILINGUALISM (if applicable)				

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

s.19(1)

s.20(1)(b)

		NO

TA Proposal [For completion by Contractor]				
9. Estimated Cost Contract <Insert additional rows as required>				
Category (Level) and Name of Proposed Resource	**PWGSC Security File Number**	**Firm Per Diem Rate**	**Estimated # of Days**	**Total cost**
[redacted]	[redacted]	[redacted]		269,197.50
Professional services estimated cost			**Total**	$269,197.50
			GST	$34,995.68
			Grand Total	$304,193.18
Travel & Living			**Estimated Cost**	
			GST	
			Total Travel & Living Cost	
			Grand Total for Labour and Travel	

TA Approval		
10. Signing Authorities		
Name, Title and Signature of Individual Authorized to Sign on Behalf of Contractor [type or print]	Contractor	Date
Francis Bilodeau Assistant Secretary, Digital Policy and Services Treasury Board of Canada Secretariat		Date
Name, Title and Signature of Individual Authorized to Sign on Behalf of PWGSC [type or print]	PWGSC	Date
11. Basis of Payment & Invoicing		
In Accordance with the article entitled "Basis of Payment" in the Contract. Payment to be made based on receipt of detailed monthly invoices for services rendered, subject to full acceptance by the Project/Technical Authority. Total of payments not to exceed the grand total. Original invoices shall be sent to the Project/Technical Authority. One copy of each invoice, together with attachments, shall be sent to the Contracting Authority.		
Financial Coding – FC: 200504 **GL:** 4664 **PA:** A002 **Fund:** B120		

GCDocs# 30978506

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

DELIVERABLE ACCEPTANCE
Department Name: _____
Task Authorization Number: _____
Contract Authority: _____
Date Submitted: [YYYY-MM-DD]
DESCRIPTION OF DELIVERABLE(S)
[Describe the deliverables accepted]
Name of the individual who accepted the deliverable: _____
Title of the individual who accepted the deliverable: _____
Date accepted: [YYYY-MM-DD]
Signature of the individual who accepted the deliverable: _____
Comments:

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

TA#2 $269,000

Statement of Work:

Four exploratory initiatives. Each initiative's objectives are to provide incremental improvements to the Departmental Recommendations.

The *Open Data Exploration, Relevance Feedback and System Logs* concepts will explore different sources of data and evaluate their impacts on the model's performance.

The *Sufficient Information* will assess whether it's possible to measure what a "good" search criteria looks like. This will follow a slightly different approach, as such will be detailed separately from the previous 3.

POC #1 Sufficient Information

The project will seek to measure when an appropriately descriptive input string has been entered and guide the user to improving his terms so that more accurate results can be returned to him. By looking at the results probability distribution, it should provide a hint as to how useful the terms used were in locating something of value. For example, terms yielding 100 results, each with a 1% probability that it is relevant to the user, would essentially translate to "*I couldn't find anything of relevance, so here's a bunch of results with equal relevance*". Conversely, if the 100 results had 1 with 95% relevance, and the other 99 were 0.05%, we might assume that the 95% relevance one is very relevant.

Explore options to indicate how performant the entered terms are so that we are confident the recommendations are as meaningful to the user as possible. If the system detects low confidence in the entered terms, the user can be informed so that he may add in additional terms to further define what they're looking for.

Methodology
Department Recommendations

- Trying to find a threshold for the sum of the probabilities of the top 5 results.
- Treat this as a binary classification problem where the top 5 is one class, and the rest are our second class
- Using the ROC curve and the F-Measure, identify the threshold that would maximize our accuracy based on user input. Call this a confidence threshold, where if the sum of the probabilities of the top 5 is above this value, then be certain that the user has provided sufficient information
- If the sum (based on user input) falls below the threshold, ask the user to provide more information

Summary Search

- Threshold on the cosine similarity of user input and retrieved documents

Deliverable

- A document detailing the experiment's respective results, with a recommendation on the way forward which would include identified thresholds.

POC#2 Relevance Feedback

This initiative would look to leverage user behaviour into the model in the hopes that it would increase overall accuracy. The idea would be that a successful recommendation would lead to fewer erratic user interactions, whereas poor recommendations may have the user navigating back and forth and/or attempting many variations of similar wording.

Additionally, inspecting the end to end user flow and evaluating success or failure could further enhance the relevance feedback. For existing summaries, a user flow ending in the initiation of an informal request of an existing ATIP report would be considered positive. For departmental recommendations, the end to end flow of a user navigating to a recommended department, submitting an ATIP request, and having that ATIP request be properly provisioned would also be considered a positive signal.

The focus for this concept will be to review erratic user behaviour and evaluate whether this leads to improved performance.

Methodology

- Relevance feedback is the process of search parameter adjustment following the initial search result retrieval. In this process, users' behaviour following the initial retrieval is analyzed to determine the relevance of the results to the query (e.g. clicking the back button, looking at multiple results, etc.). In many cases, this data can be leveraged to better optimize the engine.
- Conduct an exploratory proof of concept around the hypothesis that capturing simple user behaviour on the UI layer can create data for the Relevance Feedback model which can then be leveraged to generate more accurate results.

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Deliverable

- The results will be validated through a series of simulations and will be provided in a summary report. This report will also contain high level instructions on recommended next steps to implement the improvement.

POC#3 Relevance Feedback - System Logs

If incorporating relevance feedback shows promising results, it's possible the system logs will be a good place to look for historical traces of user flows demonstrating positive or negative signals. Assuming these logs exist and IMTD can retrieve logs for many years of activity, look to re-create a structured view of individual user workflows from these system logs. These flows could then be used in the relevance feedback model to further improve the overall system performance and accuracy.

This exploratory proof of concept will test the hypothesis that system logs (assuming they are not Protected B) could be used to reinforce positive or negative signals.

Methodology

- Assuming system logs will be coming from multiple different sources, it won't be possible to perfectly match user flows between log systems. A fuzzy matching strategy will be developed to gather up likely user flows together, using whatever data is available from the logs. For example, timestamps and source IP's may be sufficient to create a view of the user's navigation.
- Once the flows have been modeled, match them to the datasets the relevance feedback model expects and test out its impact on the overall system performance.

Deliverable

- A document outlining the experiments and its results. If successful, detailed descriptions on the fuzzy matching strategy as well as high level implementation plan to bring this strategy to production.

POC #4 Open Data Exploration

There is a wealth of data on the OpenGov site. Skim through the datasets looking for any potential data that could be used to increase the overall accuracy and robustness of the predictive models.

In parallel, a visualization team will identify a small handful of interesting datasets and explore them through data visualization. Ultimately, looking for either an informative data story or some potential for future functionality or concepts that could be added to the TBS AI Platform.

Deliverable

- Depending on what analysts find, either deliver a report recommending the inclusion of certain datasets into the models for improved accuracy or deliver a data visualization highlighting an informative data story which could lead to future platform improvements.

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Lauzon, Dominique (he/him, il)

From: Trew, Debbie
Sent: November 24, 2023 2:11 PM
To: Thompson, Ross; Annett, Megan
Cc: Herbert, Wesley
Subject: Information regarding Recommendation Engine
Attachments: Treasury Board ATIP Phase 3 - GC Strategies.pdf; ATIPclickModels_postSubmit.pdf; Lixar Tasks and Deliverables for expedis display and feedback.docx; Using artificial intelligence - Tech Corner.docx

UNCLASSIFIED / NON CLASSIFIÉ

Hello
Please find noted below and attached information that Wes and I were able to explain the Recommendation Engine. I will continue to look.

==========
==========
This is what Lixar sent us……..

tf-idf - https://en.wikipedia.org/wiki/Tf%E2%80%93idf

cosine similarity - https://en.wikipedia.org/wiki/Cosine_similarity

==========
==========
Lixar, we can provide the following preferred schema for all Search and Click Logs.

Preferred Formats: *CSV, JSON, TXT*
Schema:
session_id *Unique identifier for the user session*
datetime *Timestamp for the log*
action_type *"Q" for Query, "C" for Click*
search_term *Query search term used*
click_id *ID of a selected department (applicable when "Click")*
department_ids *An Array of Department UMD IDs; returned by the Lixar API for searches, ordered by Rank (1,2,3, etc)*

I hope this helps, if you would like us to jump on a call to talk around this more, please let us know.

Therefore, we will look to begin development against this schema, however, if this proves difficult or you require an alternative, please let us know as soon as possible.
==========
==========

Sites and files crawled for the Recommendation Engine -
) the summaries files

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

2) departmental reports --> departments listed on this wesbite https://www.canada.ca/en/treasury-board-secretariat/services/departmental-performance-reports/list-institutions-2016-17-departmental-results-reports.html
3) combing through various gov websites (exact list of url's would take some time to dig out)
4) Institution ATIP Pages.xlsx --> a file shared by the client at the beginning of atip phase 2.
canada.ca
2016-17 Part III - Departmental Results Reports (DRR) - Canada.ca
TBS is providing this listing to make it easier for you to access the Departmental Performance Reports on various government institution websites. Please note that some links might not work until institutions post their reports on their websites.

Here is an example of what the Lixar API returns to us:

{

"StatResults":[

{

"Term":"tax",

"Stats":{

"Last30Days":"0",

"Last90Days":"1",

"MostCommonDepartment":{

"Name":"Canada Revenue Agency | Agence du revenu du Canada",

"UmdNumber":"72"

}

}

},

{

"Term":"money",

"Stats":null

},

{

"Term":"canada",

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

"Stats":{

"Last30Days":"0",

"Last90Days":"1",

"MostCommonDepartment":{

"Name":"Treasury Board of Canada Secretariat | SecrÃ©tariat du Conseil du TrÃ©sor du Canada",

"UmdNumber":"117"

}

}

}

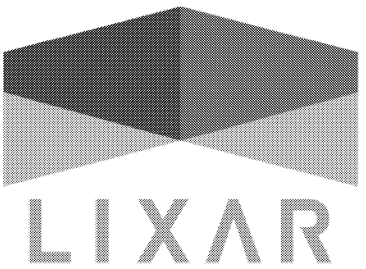
]

}

Debbie Trew
Senior Project Officer
Open Government, Office of the Chief Information Officer
Treasury Board of Canada Secretariat / Government of Canada
Debbie.Trew@tbs-sct.gc.ca / 613-219-9626

Responsable principal de projet
Division du gouvernement ouvert, Bureau du dirigeant principal de l'information
Secrétariat du Conseil du Trésor du Canada / Gouvernement du Canada
Debbie.Trew@tbs-sct.gc.ca / 613-219-9626

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*



gcstrategies

Statement of Work

ATIP Phase 3
Treasury Board Secretariat

September 14, 2018

LIXAR.COM | INFO@LIXAR.COM | @LIXARIT
Mobile . IoT . Data Science . Cloud

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Copyright © 2018 Lixar I.T. Inc.

All rights reserved. This document is intended for the person(s) and/or organization(s) named on the title page of this document. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, or other electronic or mechanical methods, to other parties, without the prior written permission of Lixar I.T. Inc., except in the case of brief quotations embodied in critical reviews and certain other noncommercial uses permitted by copyright law. For permission requests, write to Lixar I.T. Inc. at the address below.

Lixar I.T. Inc.

373 Coventry Road

Ottawa, ON K1K 2C5 CANADA

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

CONTENT

Summary 4

About Lixar 4

Proposed Engagement 5

Summary 5

Productionize 5

Sufficient Information 5

System Log Analysis & Relevance Feedback 6

Extend 7

Ontology Development Phase 1 7

Experiment 10

Semantic Search POC 10

Topic Modeling POC 10

Data Augmentation POC 11

Project Constraints 12

Dependencies 12

Risks 12

Timeline/Milestones 14

Normalization/Warranty Terms 14

Changes and Additions 15

Termination 15

Expiry 15

Pricing and Billing 16

Cost Estimates 16

Billing Schedule 16

Error! Bookmark not defined.

000057

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Summary

Project Title	ATIP Phase 3
Assessment Date	September 14, 2018
Target Delivery Date	TBD
Assessment Version	1.0
Statement of Work ID(s)	RFADN-113. This ID is referred to as "SOW" throughout this document.
Based on	Based on a series of discussions between the Treasury Board Secretariat, GC Strategies, and Lixar

About Lixar

Lixar is a technology company that provides end-to-end technology solutions for enterprise. We stay on the leading edge of technology by embracing a culture of innovation and welcoming change. Through our intense focus on research and development, our data & analytics and cloud practices, and our ability to identify emerging technology, Lixar stays true to our purpose - to create, innovate, and inspire. Lixar's technology ecosystem consists of IoT, Software & Mobile, Data Science, and Cloud.

Our data science practice, founded in 2011, is built around a team of 40 data scientists, data engineers, data visualization analysts. We create novel algorithms to optimize, automate, and predict by leveraging machine learning. Our data practice has successfully delivered programs for Delphi, General Motors, NASCAR, Jamieson Vitamins, Air Canada, and a number of other enterprise clients.

Our company has grown organically over since 1999 to over 200 employees in multiple office locations.

Objective

The overall objective of this phase of the ATIP project is to continue to enhance the overall accuracy and utility of both the Search and Recommendation systems. Phase 3 proposes that we focus on the productionization of previous POCs to realize the improvements in accuracy that were discovered, develop new features to further increase the utility of Search, and engage in new POCs to refine and augment the Recommendation engine in future phases.

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Proposed Engagement

Summary

- Productionize
  - Sufficient Information
  - System Log Analysis & Relevance Feedback
- Extend
  - Ontology Development Phase 1
- Experiment
  - Semantic Search POC
  - Topic Modeling POC
  - Data Augmentation POC

Productionize

Sufficient Information

Building upon the methodology of determining the sufficient information threshold which is outlined in the PoC, it is possible to advice users to modify their request such that the system would return them the target department with higher accuracy.

Productionazing the sufficient information can be done in two levels:

Updating the sufficient information threshold

As more summaries get added to the list of ATIP summaries and the summaries get augmented using other data sources, the sufficient information threshold is subjected to change. It is however, possible, to automate the process of obtaining the threshold based on the Precision-Recall curve and the F-Score curve. The frequency of re-calculating the sufficient information threshold will be experimentally determined. The steps are listed below:

Step 1: Calculate to top-5 likeliest departments sum probability
Step 2: Obtain TP, FP, FN
Step 3: Obtain Precision, Recall, and F-Score
Step 4: Measure optimal F-score and the corresponding sum probability

Guiding the users to improve their search entries

000059

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Once we obtain the threshold it is possible to guide the user to change his/her search queries to guarantee better outcome based on sufficient information provided. This process is completed following these steps:

Step 1: The user enters a request
Step 2: The model predicts the top 5 likeliest department and their probabilities
Step 3: If the sum probability is above threshold:

- The model returns results to the user

If the sum probability is below threshold:

- The system notify the user to change the search terms due to insufficient information

System Log Analysis & Relevance Feedback

Building on the success of the POCs in the previous phase, Lixar is recommending the productionization of both web log analysis and relevance feedback from user click data. As demonstrated in the POCs, this data provides implicit user feedback that can be mined to fine-tune the relevance scores returned by the Department Recommendation AI system. **Integrating relevance feedback into the ATIP system would require significant involvement from the IMTD team as we would be reliant on them capturing and submitting data.**

The simplest measurable way that users interact with search results is by clicking on them. Probabilistic models of click-based user interactions with search results are therefore called 'click models'. Click models define user behaviour in terms of events that are either *observable* or *hidden*. Whether or not a particular search result is relevant is hidden; users don't directly label the relevance of search results. What isn't hidden are the observable events, in this case the user clicks. So while relevance isn't immediately observable, click models assume that it depends on the sequence of observable clicks. An appropriate model of user behaviour can therefore be used to assess the relevance of search query results.

Relevance adjustment for the ATIP system will be calculated using a Dynamic Bayesian Network click model. This model captures user behaviour by considering four 'events' that take place during user interaction with search results:

- *E – Examination* of a department (the user reads the department name and caption).
- *C – Clicking* on a department.
- *A* – The *attraction* of the user to a department and its description. This is the **perceived relevance** of the department.
- *S* – The *satisfaction* of a user department after it has been clicked. This is the **actual relevance** of the department.

The main advantage of this click model compared to many others is that it distinguishes between perceived relevance and actual relevance. In the case of the ATIP system, the department with actual relevance is the one where an ATIP request is submitted.

000060

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Over time, as users interact with the recommendation system, data will accumulate for many unique queries, each of which are given a unique query id. To improve model performance, queries that are not strictly equivalent but are in essence the same will be matched. For example, the queries "the Canadian government" and "government of Canada" are for the same topic, but would be assigned different query ids if strict equivalence was required. A query matching framework was therefore constructed that uses the already developed backbone of the pre-processing procedure for the Department Recommendation AI system. In this procedure, each row of the Document Term Matrix (DTM) is a Term Frequency-Inverse Document Frequency (TF-IDF) vector representation of a given ATIP query, where weights are assigned to terms based on their frequency of appearance within individual documents (TF) relative to the corpus as a whole (IDF). Terms that appear uniquely in queries for specific departments therefore have a higher weight than terms that appear in many departments. Using this approach, queries that contain similar terms will have similar vector representations. When the similarity crosses an appropriate threshold, queries are considered to be the same.

The output from the DBN click model can be easily integrated with the current recommendation system to refine and improve the search results provided by the system, which will ultimately lead to increased likelihood of ATIP requests being submitted to the correct department. To combine the relevance scores produced by the Department Recommendation AI system and the click model, an overall relevance score that combines the results from both will be used. Initially, when there are no records in the system logs for a given query, the overall relevance will depend exclusively on the scores produced by the Department Recommendation AI system. As user clicks on query results accumulate, the relevance scores generated by the click model will become more statistically reliable. As such, the adjustments made to the overall relevance scores for a given query will be weighted more heavily to click model analysis as user interactions for a given query accumulate in the system logs.

Extend

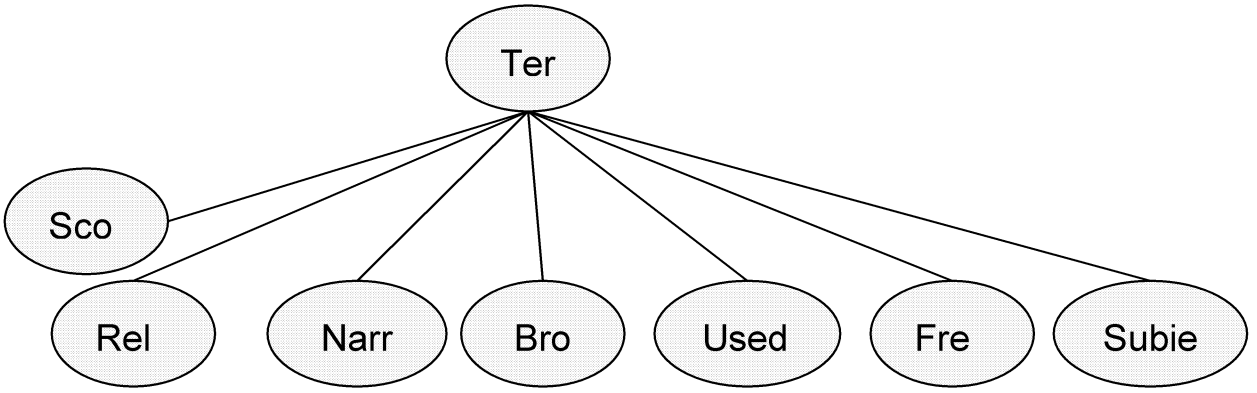
Ontology Development Phase 1

An ontology can assist with the identification of key terms and the relation between them, in order to build an expert system. The expert system can then augment the existing classifier and the search engine. The proposed work highlighted below is not a proof of concept, but instead marks the first phase of an ATIP ontology that will provide the foundation for future evolution. A properly crafted ontology has huge impacts on natural language processing and will help the system understand relevant domain language at a much deeper level.

One approach to integrate ontology into the search engine/classification is to create domain ontology. Domain ontology in the context of ATIP, would refer to Government of Canada and its related agencies/departments. We list out all the concepts related to the Government of Canada along with the relations. One way of building a lightweight domain-specific ontology is to use the taxonomy of Government of Canada. The taxonomy includes around 4900 terms and their related information such

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

as a French translation of a term, its related terms, related subject category, what the term is used for, and the broader and narrower term if exists.

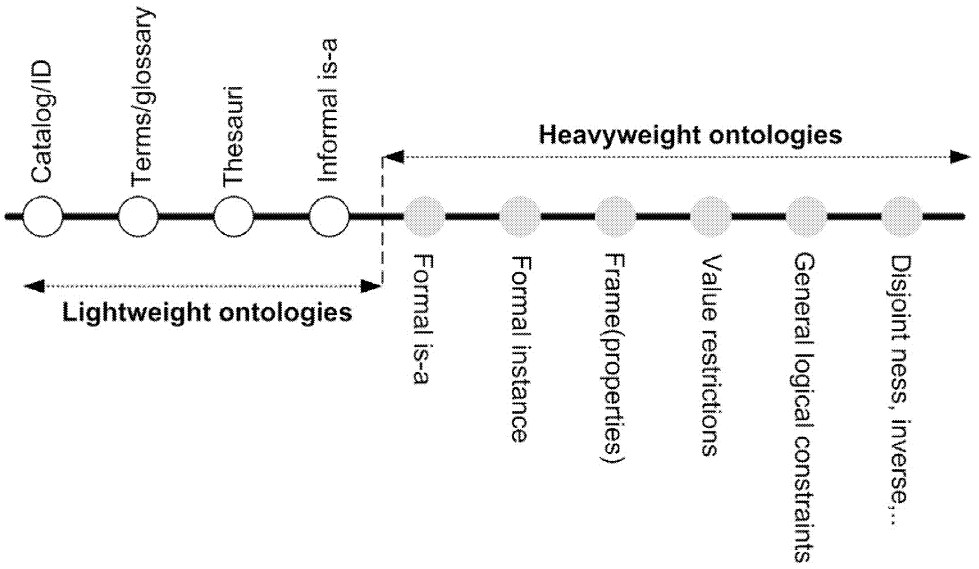


For example, the term "Health" and "Livestock" are categorized as follows:

Term	Category	Category Value
Health	French	Soins de santé
	Related Term	Caregivers, Health, Health care facilities, Health care industry, Health care products, Health care quality, Health care system, Medicine, Telehealth, Therapy
	Subject Category	HE, Health and Safety
	Used For	Home care, Medical care, Nursing, Psychiatric care
	Broader Term	N/A
	Narrower Term	N/A
	Scope Notes	N/A
Livestock	French	Bétail
	Scope Note	Animals kept of raised for use, food, or other products.
	Used for	Farm animal, Herds

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

	Broader term	Animals
	Narrower term	Cattle, hogs, horses, sheep
	Related Term	Animal husbandry, Domestic animals, grassland, manure, meat, poultry, slaughtering
	Subject Category	AG Agriculture



An Ontology Expressiveness Spectrum adapted from McGuinness (2003)

Depending on the available information and collaboration with subject matter experts, a choice will be made to proceed with either lightweight or heavy weight ontologies.

Once the ontology is built, it could be integrated into the semantic search engine in two phases of the retrieval process:

- Disambiguation in order select the search domain (e.g. virus in computer science vs. virus in medical field)
- Relevance Feedback (by adding other keywords similar to the concept expressed in the query)

The other usage of ontology is building an ontology-based classification. The summaries are first preprocessed using existing preprocessing pipeline of tokenization, stemming, stop word removal, and punctuation removal. The summaries are split into a test set and a training set where the training set is used as an input to the ontology-based algorithm. We then apply Word Embedding (e.g., Word2Vec) to the training set in order to capture the similarities between each concept related to a specific department in the ontology and each term in the corresponding summary. In the following step, the

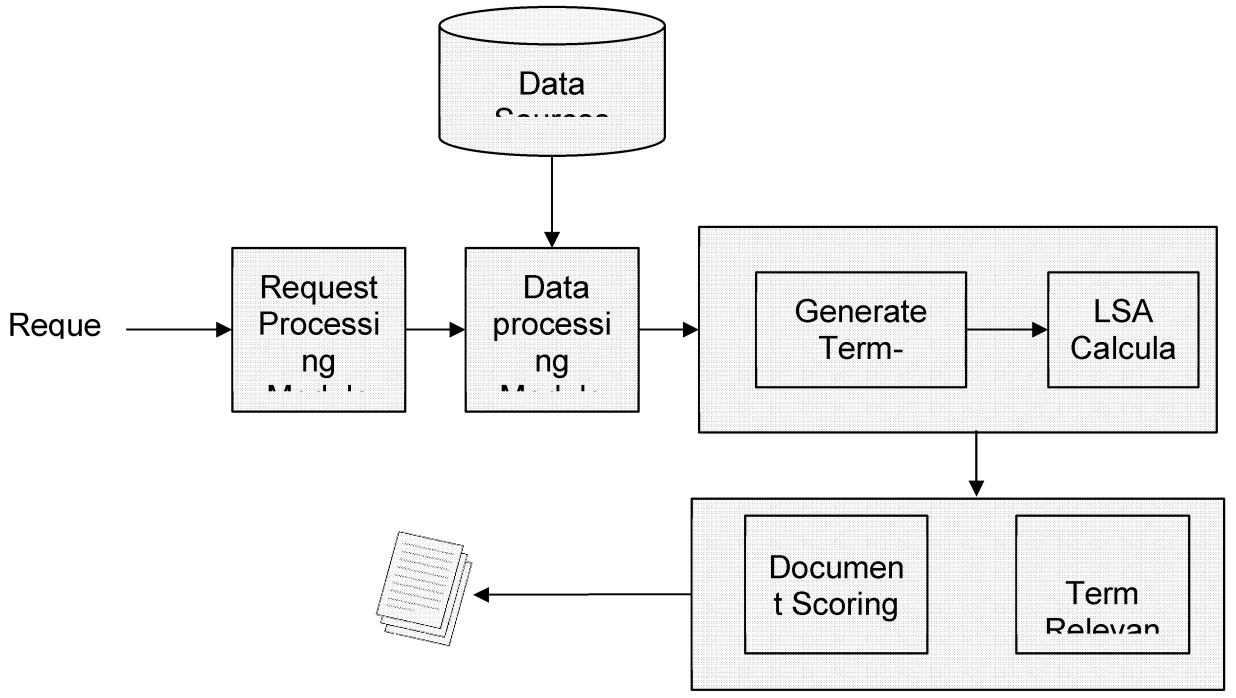
Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

similarity between terms in the test set and each concept related to a department are summed up in order to compute the total similarity between a test summary and a department. All departments with positive total similarity are selected as potential labels. The positive summaries are ranked and potential departments are obtained.

Experiment

Semantic Search POC

Latent Semantic Analysis (LSA) allows for the retrieval to be done based on the concepts available in text, rather than individual words (TF-IDF). In other words, it allows a search engine to understand synonyms. Using synonyms through a document would eventually help a search engine to understand the content.



LSA runs a matrix operation called Singular Value Decomposition (SVD) on the TermDocumentMatrix. With SVD, we factor the matrix into a simpler form. This is followed by LSA calculation which shows, mathematically, that our matrix can be decomposed around the corpus's main ideas. With LSA, we can break up a largeTDM into semantically meaningful parts.

Topic Modeling POC

By identifying a set of topics, it's possible to divide the summaries into 2 or more categories, with each category containing records that share similar concepts. Following this, a classifier can be built for each

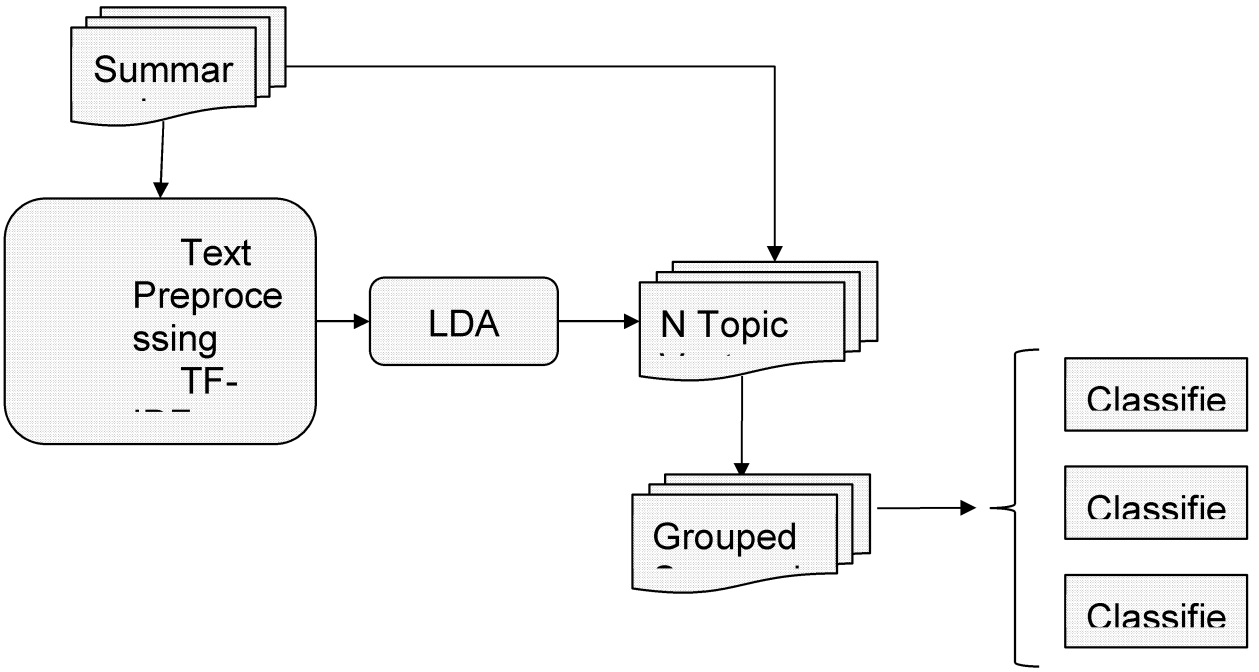
000064

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

of these categories, potentially resulting in the reduction of the bias of the classification module without increasing the variance.

There are several algorithms that can be used for topic modeling. Some of the main algorithms are LDA (Latent Dirichlet Allocation), LSA or LSI (Latent Semantic Analysis or Latent Semantic Indexing) which uses SVD on the Document-Term Matrix based on Linear Algebra, and NMF (Non-Negative Matrix Factorization) based on Linear Algebra.

The overall system design is shown in the following figure:



Following text pre-processing steps, TF-IDF representation of the summaries is obtained. LDA model is trained and saved. After identifying N topics, the summaries associated with each topic are grouped together. Each group will be used in the next level to build a classifier. The choice of classifier can be experimentally validated. When a new request is send to the system, using the saved LDA model, it will categorize the request into one of the topics. Once the topic is determined the classifier associated with that topic is triggered and is used to predict the department.

Data Augmentation POC

Data augmentation has been used in various image analysis works, though it can also be applied to texts. The purpose of this piece is to artificially produce new records that can be used to increase the accuracy of the classifier.

The text augmentation is a an active research area currently. We will investigate the benefit of text augmentation in the context of ATIP summaries using the summaries and/or a combination of summaries and their respective departments.

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Data Augmentation based on summaries: In this approach, we replace words in the summaries (following the pre-processing steps including tokenization, stop word removal, and punctuation removal) with their closet synonyms. We can use WordNet and Word2Vec and apply different closeness measure to determine those synonyms. The extracted synonyms will be then augmented into the text.

Data Augmentation based on departments' titles: Another approach is to augment the requests by adding, repeating keyword terms based on the department titles the requests are associated with. The keywords in the title and their closest synonyms will be used to augment the summaries.

Depending on the results, we may use either approach or both combined to enhance our performance.

Project Constraints

Dependencies

The following list outlines the dependencies for the project:

- If TBS selects the Log Analysis and User Feedback POC for productionization, heavy involvement will be needed from the IMTD/OpenGov development team to capture data and submit that data to our system.

Note: In order to meet established project timelines, external dependencies need to be completed by the mutually agreed upon due date(s).

Risks

The following risks have been identified:

- As with all POCs, the outcome is uncertain. While we have a high degree of confidence that the suggested POCs will add significant value to the overall system, there is always the risk of any experiment yielding no net positive results.
- Leveraging on the previous risk, some of the PoCs including Topic Modeling, Semantic Search, and Sufficient Information will depend heavily on the quality of data in the summaries. Lixar anticipates two possible mitigation strategies. First, as more data gets added to the pool of available summaries, better results could be obtained. The PoCs will be using the most up to date summaries, however, since the summaries get updated frequently, the same implementation could lead better results. Second, some of the PoCs presented in this proposal, including "building an ontology" and "augmentation" aim to enhance the quality of the data. Consequently, this could improve the results of other PoCs as well.

000066

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Timeline/Milestones

The following timeline is provided as a guideline. Actual timelines will be discussed and agreed to by both parties upon project sign-off:

Milestone	Delivery Week
Project Signoff	Week 0
Productionize Sufficient Information	Week 6
Semantic Search POC	Week 8
Productionize Log Analysis & User Feedback	Week 12
Topic Modeling POC	Week 12
Data Augmentation POC	Week 16
Ontology Development POC	Week 20

Normalization/Warranty Terms

Lixar shall correct any material deficiencies in any software(s) for a period of thirty (30) days following deployment into a production environment ("normalization period"). Such work will be scheduled to be resolved within a commercially reasonable period of time during normal business hours. Lixar shall have no obligation to correct deficiencies to the extent caused by (but not limited to):

1. Third-party software or services relating to the software;
2. Modifications made to the software by any party other than Lixar;
3. Use of the software in a manner for which it was not designed, including its use in connection with a technical environment other than as agreed upon.

Following the normalization period, the Treasury Board Secretariat shall be solely responsible for the management, maintenance, and support of the software. If required, such services can be separately contracted with Lixar. Lixar offers fixed annual rate or monthly time and materials-based agreements to facilitate post-normalization technical support and defect corrections.

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Changes and Additions

Any net new additions in scope must be submitted to the Lixar Project Manager. This includes (but is not limited to) change requests, project extensions due to delays and/or validating any Dependencies. Such additions may impact project timelines or project costs and thus require supplementary assessment.

Termination

Should a decision be taken by the Treasury Board Secretariat to terminate this initiative, a written notification with reasonable notice of at least one month will be issued to the Lixar Account Executive. The Treasury Board Secretariat shall pay Lixar for all services as specified in the Agreement and expenses, if any, up to the effective date of the termination, including work in progress. All work in progress shall be deemed accepted as is.

Expiry

The information presented in this Statement of Work is valid for a period of 90 calendar days from the Statement of Work submission date. If the expiry date falls on a non-business day for either Lixar or the Treasury Board Secretariat (e.g., weekend, holiday), the expiry date is considered to be the next business day.

The Treasury Board Secretariat must approve and sign off on the Statement of Work and return it to Lixar prior to the end of business (5:00pm EST) on the expiry date. Otherwise, the information presented in this Statement of Work becomes null and void.

000069

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Pricing and Billing

The cost estimates and billing schedule represent the contract between **GC Strategies** and Lixar to perform the work for the Treasury Board Secretariat as detailed above.

Cost Estimates

The following cost estimates have been defined:

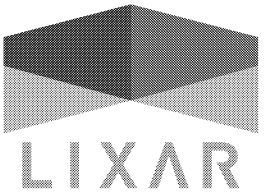
Milestone	Price
Productionize Sufficient Information	
Productionize Log Analysis & User Feedback	
Semantic Search POC	
Ontology Development POC	
Topic Modeling POC	
Data Augmentation POC	

Notes:

- All pricing excludes any value add tax (e.g., HST) where applicable.
- All pricing in CAD.

Billing Schedule

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*



Government of Canada | Gouvernement du Canada

Access to Information and Privacy (ATIP)

Department Recommendations: Relevance Adjustments Using System Logs

Proof of Concepts: *Relevance Feedback* and *System Logs*

September 13, 2018

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Copyright © 2018 Lixar I.T. Inc.

All rights reserved. This document is intended for the person(s) and/or organization(s) named on the title page of this document. No part of this publication may be reproduced, distributed, or transmitted in any form or by any means, including photocopying, recording, or other electronic or mechanical methods, to other parties, without the prior written permission of Lixar I.T. Inc., except in the case of brief quotations embodied in critical reviews and certain other noncommercial uses permitted by copyright law. For permission requests, write to Lixar I.T. Inc. at the address below.

Lixar I.T. Inc.
373 Coventry Road
Ottawa, ON K1K 2C5 CANADA

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Contents

1 Introduction **4**

2 The ATIP Click Model **4**
2.1 Click Model Foundations 5
2.1.1 Description of Random Variables 5
2.1.2 Model Parameters 6
2.1.3 Relationships Between Variables 7
2.2 The Model: Dynamic Bayesian Network 7

3 Model Input: System Click Logs **9**

4 Query Matching **10**
4.1 Calculating Similarity 10
4.2 Matching Examples 11

5 Model Output: Relevance Probabilities **13**

6 System Integration: Relevance Adjustments **15**
6.1 Relevance Adjustment Calculation 15
6.2 Example Adjustment 16

7 Conclusions **17**

000073

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

1 Introduction

The Access to Information Act governs access to information contained in federal government records. Canadian citizens, permanent residents, persons present in Canada, or corporations can request records under the control of Public Safety Canada. In order to obtain those records, a requester may complete the Access to Information Request Form or send a letter to the ATIP Office referencing the Act and describing the information sought. Requests for information initiated by other government institutions, including portfolio agencies, require that the requester refers to the list of departmental Access to Information and Privacy Coordinators. It is, however, a cumbersome task to determine the department to which a request should be directed.

Working towards a solution to this problem, Lixar have developed a Department Recommendation AI system [1]. The system receives a text search query as input and applies classification and machine learning techniques to find the set of N departments with the highest 'relevance' scores calculated by the recommendation system. In the user interface, these departments are then listed in descending order. In this report we investigate how implicit user feedback stored in system logs can be used to fine-tune the relevance scores according to their *actual* relevance. Actual relevance for the ATIP system will be calculated using a Dynamic Bayesian Network click model, which will be introduced in Section 2. The information that the model requires from system logs of user clicks is described in Section 3. Section 4 is concerned with how user queries in system logs will be matched to improve model performance. This is followed by a description of output from the Dynamic Bayesian network click model in Section 5, and how it will be integrated in the output from the Department Recommendation AI system in Section 6. Conclusions are made in Section 7.

2 The ATIP Click Model

One way to assess the relevance of search query results is to have them analyzed and interpreted by a subject matter expert. For a given query, such experts could provide a label that would directly identify the relevance of each 'document' in the results. Here, a document refers to any specific query result from a web-based search system, such as a website, a video, or in the case of the ATIP system, a department. In lieu of expert labels, user interactions with web-based search results provides implicit feedback about

000074

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

their relevance.

The simplest measurable way that users interact with search results is by clicking on them. Probabilistic models of click-based user interactions with search results are therefore called 'click models'. The specific click model that Lixar have developed for use with the ATIP department recommendation system is described in this Section. A general introduction to click models is given in Section 2.1, followed by a description of the specific click model that will be used for the ATIP system in Section 2.2.

2.1 Click Model Foundations

The term click model was first introduced in 2008 [2], and since that time the development of click models has been ongoing, with research contributions from both industry (such as Microsoft [3], Yandex [4], and Yahoo [5]) and academia. While there are various versions of click models, many models are constructed using the same basic foundations. Here we describe these foundations and introduce the relevant terminology.

In general, to define a click model unambiguously requires: a description of its random variables (2.1.1), a description of the dependence of the variables on model parameters (2.1.2), and definition of the relationships between variables (2.1.3.).

2.1.1 Description of Random Variables

Click models define user behaviour in terms of events that are either *observable* or *hidden*. Whether or not a particular search result is relevant is hidden – users don't directly label the relevance of search results. What isn't hidden are the observable events, in this case the user clicks. So while relevance isn't immediately observable, click models assume that it depends on the sequence of observable clicks.

Each event for click models is described with a binary random variable X: if the event doesn't happen $X = 0$, if it does $X = 1$. We will assume that for the ATIP system the Search Engine Results Page (SERP) will have $N = 5$ departments. We will also assume that the department name is shown on the SERP along with a caption that briefly describes the department. The events that will be considered during user interaction with the SERP are:

- E – *Examination* of a department (the user reads the department name and caption).
- C – *Clicking* on a department.

000075

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

- A – *Attractiveness* of a department and its description. This is the **perceived relevance** of the department.
- S – *Satisfaction* of a department after it has been clicked. This is the **actual relevance** of the department.

In addition to events, other notation used for the model is d_r, where d refers to a specific department from the set of departments returned on the SERP, and r refers to the rank of a department in the set. Department rank is determined by the relevance scores returned by the Department Recommendation AI system.

2.1.2 Model Parameters

The 'random click model' serves as a simple example for introducing the concept of model parameters, as it has only one:

$$P(C_d = 1) = \rho. \tag{1}$$

The parameter ρ is the probability that department d will be clicked. For this model, all departments have the same probability of being clicked. Research clearly indicates, however, that users don't randomly click on search engine results randomly with equal probability [6]. That is, there is a *position bias* in the documents that are clicked, where users are much more likely to click on documents presented at the top of the list.

To incorporate user preference for clicking on higher ranked search results, rank dependence can be introduced to the model so that

$$P(C_r = 1) = \rho_r. \tag{2}$$

While this model (called the 'rank-based click through model') is still very simple, it can be used as a baseline to compare the performance of more advanced models, as it was in Ref. [7].

If we assume that user interaction with the system depends on both the departments returned by a query and the query itself, then parameters are required for each query-department pair, which can be written as

$$P(C_d = 1) = \rho_{dq}. \tag{3}$$

This is called the 'document-based click through model', which was first introduced in Ref. [2].

While the position bias is captured in Eq. 2, and document bias is captured in Eq. 3, these simple models do not capture the hidden events that

000076

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

occur during user interaction with a SERP (such as examination, attractiveness, and satisfaction). To do so requires parameterizing other aspects of user behaviour (as described in Section 2.2), and establishing relationships between the variables (which can also be parameterized).

2.1.3 Relationships Between Variables

The 'examination hypothesis' is included in many click models, and is an example of how relationships between variables are defined in click models. It establishes a relationship between clicking, examination, and attractiveness to be

$$P(C_d = 1) = P(E_d = 1) \cdot P(A_d = 1). \tag{4}$$

This states that a user will click a document if, and only if, they both examined the document and were attracted by it (they perceived it to be relevant). For the ATIP search engine results, attractiveness is a characteristic of the department name and its caption.

2.2 The Model: Dynamic Bayesian Network

In general, click models define dependencies between events, where the goal is to estimate full or conditional probabilities of the corresponding random variables. For example, for ATIP search engine results, $P(S = 1)$ represents the full probability that the user is satisfied after clicking on a department (meaning they will submit a request to that department), and $P(C = 1|E = 1)$ represents the conditional probability that a user will click on a department given that they examined its name and description.

In this Section we define the model parameters and relationships between variables for the ATIP click model, which is a Dynamic Bayesian Network (DBN) [5]. The main advantage of this click model compared to many others is that it distinguishes between perceived relevance and actual relevance. In the case of the ATIP system, the department with actual relevance is the one where an ATIP request is submitted.

Using the random variables E, C, A, and S as described in Section 2.1.1, the DBN click model is completely described by the following set of proba-

000077

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

bilities:

$$P(C_r = 1) = P(E_r = 1) \cdot P(A_r = 1) \tag{5}$$
$$P(A_r = 1) = \alpha_{qd_r} \tag{6}$$
$$P(E_1 = 1) = 1 \tag{7}$$
$$P(E_r = 1 \mid E_{r-1} = 0) = 0 \tag{8}$$
$$P(S_r = 1 \mid C_r = 1) = \sigma_{qd_r} \tag{9}$$
$$P(E_r = 1 \mid S_{r-1} = 1) = 0 \tag{10}$$
$$P(E_r = 1 \mid E_{r-1} = 1, S_{r-1} = 1) = \gamma \tag{11}$$

Note that in the above equations, for simplicity the rank r for the random variables A, E and C refers to the rank of a specific department d, even though d isn't subscripted. An explanation of each equation in the DBN model follows:

- Eq. (5) – This is the examination hypothesis presented in Section 2.1.3.
- Eq. (6) – Perceived relevance of a department is a model parameter that depends on the department-query pair.
- Eq. (7) – It is assumed that the user will examine the top ranked department.
- Eq. (8) – There is zero probability that a user will examine a department without first examining the departments at higher ranks.
- Eq. (9) – σ_{dq} is the actual relevance of a department to a given query, where satisfaction requires that a user clicks on the department.
- Eq. (10) – The probability of a user examining a document at a lower rank after being satisfied with one at a higher rank is zero. That is, once satisfied, the user will not continue to click on other departments.
- Eq. (11) – γ is the continuation probability for a user that clicked on a department but was not satisfied by it. $1-\gamma$ is therefore the probability that a user will abandon their query without being satisfied by it. For the ATIP model we assume that ($\gamma = 1$); users will not abandon their session until they have found the appropriate department.

A graphical model of the DBN described by Eqs. 5-11 is shown in Fig. 1. The Bayesian network in each box represents user interaction with a specific department returned by the search engine. The box on the left is for user

000078

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

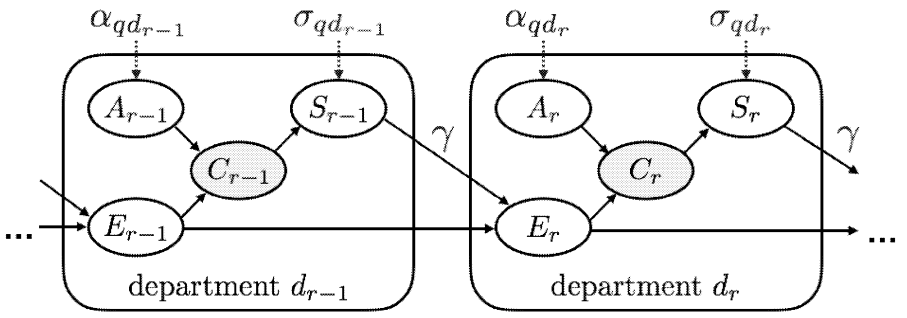


Figure 1: The DBN displayed as a graphical model. Random variables are shown as ovals, with hidden variables shown as white and observed variables shown as yellow. Solid arrows between circles show the direction of probabilistic dependence. Model parameters are shown in blue.

interaction with the department at rank $r-1$. Whether or not the user clicks (C_{r-1}) on the department depends if they examined the department (E_{r-1}) and perceived that it was relevant (A_{r-1}). The probability of the user perceiving relevance is $P(A_{r-1}=1)=\alpha_{qd_{r-1}}$. The probability of the department having actual relevance, $P(S_{r-1}=1)=\sigma_{qd_{r-1}}$, depends on whether or not the user clicked on the department. If the user is not satisfied with the department at rank $r-1$, the probability of them examining the next department d_r is parameterized by γ. For each user session with the ATIP search system, this sequential flow from one department to the next will be captured by the system click logs.

3 Model Input: System Click Logs

All of the information required for DBN model calculations is contained in the example click log shown in Fig. 2. Each column contains:

- session_id – This number uniquely identifies each session that a user has with the system. The session id must be the same for the initial query and the clicks on its results.
- action_type – Defines whether or not the user action was a click ('C') or a query ('Q').
- query_click_id – Gives the query id if the user action was a query. If the query already exists in the database, it will be given the appropriate existing id, if it doesn't, it will be given a new id. A description of how

000079

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

uniqueness of a query is determined is given in Section 4. If the action type was a click rather than a query, this field will instead contain the department id.

- dept_id_r – When the user action is a query, the departments for each rank r are given in these fields. For click actions these fields are blank.

session_id	action_type	query_click_id	dept_id_1	dept_id_2	dept_id_3	dept_id_4	dept_id_5
0	Q	3	104	111	108	112	113
0	C	111					
1	Q	3	104	111	108	112	113
1	C	104					
1	C	111					
1	C	108					
2	Q	3	104	111	108	112	113
2	C	111					
3	Q	3	104	111	108	112	113
3	C	104					
3	C	111					
3	Q	3	104	111	108	112	113
3	C	104					
3	C	111					
3	C	108					
3	C	109					

Figure 2: An example click log required by the DBN model. Four example user sessions with the same query (#3) are shown.

4 Query Matching

While two queries may not be strictly equivalent, they could in essence be the same. For example, the queries "the Canadian government" and "government of Canada" are for the same topic, but would have two different query ids if strict equivalence was required for query matching. In this Section we address the need to establish a query matching framework so that very similar queries can be assigned the same query id. This will improve system performance by reducing the number of unique queries, which will lead to DBN model output with increased statistical significance.

4.1 Calculating Similarity

The output of the pre-processing procedure for the Department Recommendation AI system already developed is a document-term-matrix (DTM). Each row of the DTM is a TF-IDF (Term Frequency-Inverse Document Frequency) vector representation of a given ATIP query, where weights are assigned to terms based on their frequency of appearance within individual

000080

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

documents (TF) relative to the corpus as a whole (IDF). Terms that appear uniquely in queries for specific departments therefore have a higher weight than terms that appear in many departments. Using this approach, queries that contain similar terms will have similar vector representations.

The similarity of vector representations for any two given queries can be calculated using cosine similarity, which is commonly used in text mining and information retrieval where high-dimensional feature spaces are common. This measure computes the cosine of the angle between two query vectors $\mathbf{A}$ and $\mathbf{B}$ as

$$\text{similarity} = \cos\theta = \frac{\mathbf{A} \cdot \mathbf{B}}{||\mathbf{A}||\ ||\mathbf{B}||}. \tag{12}$$

The maximum magnitude for similarity is therefore similarity= 1, while the minimum is similarity= 0. Using this measure, a threshold for uniqueness can therefore be established that would allow sufficiently similar queries to be matched and given the same query id.

4.2 Matching Examples

For a test database with 93 syntactically unique queries, Fig. 3 shows how the number of unique queries decreases as the threshold is lowered. At high thresholds just below one, where the criteria for uniqueness is most stringent, no queries matched, and the number of unique queries stays the same. As the threshold is lowered the matching criteria relaxes, and the number of unique queries decreases. This trend continues down to the smallest thresholds, where the number of unique queries plateaus at 49. At the plateau, 44 of the original queries have been matched with another query, and the 49 remaining queries do not share a single term from the corpus.

As a specific example, consider the following five queries:

q_1: Contracts for the Canadian warships to be built in Halifax

q_2: Contracts for the Canadian warships

q_3: Canadian warship contract Halifax

q_4: Canadian warship contract

q_5: Canadian warship

The cosine similarities calculated for these queries are given in Table 1. The diagonal starting from the top left shows that the cosine similarity of a query with itself (similarity=1). Queries q_2 and q_4 also have similarity=1,

000081

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

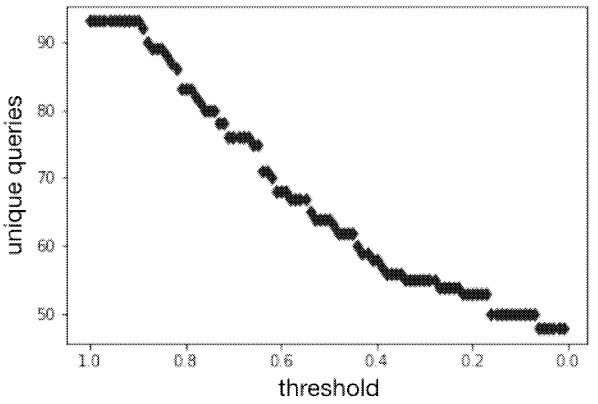


Figure 3: The number of unique queries changes depending on the threshold value for cosine similarity.

since after pre-processing of the queries the 'stop words' are removed ('for' and 'the') and plurals ('contracts') and considered the same as their singular form. The queries with the next highest similarity are q_1 and q_3, which share three words with relatively high TF-IDF scores ('warship', 'contract', and 'Halifax'). 'Canadian', on the other hand, has a low TF-IDF score because it's contained in queries for many different departments. The queries with the lowest similarity scores are q_1 and q_5, with similarity=0.288. Being the most specific query of the sample set, q_1 contains the most information, whereas q_5 has the least, leading to their lack of similarity. Based on this and similar experiments, we conclude that a reasonable threshold for query matching is similarity=0.6, however this is subject to change with future testing and system log analysis.

	q_1	q_2	q_3	q_4	q_5
q_1	1	0.419	0.694	0.419	0.288
q_2	0.419	1	0.603	1	0.687
q_3	0.694	0.603	1	0.603	0.415
q_4	0.419	1	0.603	1	0.687
q_5	0.288	0.687	0.415	0.687	1

Table 1: Cosine similarities for the list of example test queries.

000082

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

5 Model Output: Relevance Probabilities

In this Section we present a simple example of DBN calculation results for a click log with 10 different user sessions, all for the same query. The following five department ids are returned by the recommendation system, in order of highest rank (highest relevance score) to lowest rank:

d_1: 104
d_2: 111
d_3: 108
d_4: 112
d_5: 113

The click patterns for different user sessions, each enclosed in {...}, are as follows:

{108}
{104, 111, 108}
{104, 111}
{104, 111, 108}
{111, 108}
{104, 108}
{104, 111, 108}
{111}
{104, 111, 108, 112, 113}
{104}

Examining user clicks on the first and third ranked departments, with department ids #104 and #108 respectively, we see that six of the user sessions end with a click on department #108. In these instances the users are satisfied once they click on the department, indicating that the department has *actual relevance*. Department (#104), on the other hand, was clicked seven times. However, on six of these occasions the user proceeded to click at least one more department. This indicates that department #104 has *perceived relevance*, but not actual relevance.

The model calculations for this set of click patterns reflect these interpretations about the user logs, as shown in Fig. 4. Department #108 has high probability of perceived relevance (0.8889) and actual relevance (0.7778), since it was often clicked and it was often the last department clicked. The perceived relevance of department #104 (0.6667) is lower than #108 because it was examined but not clicked on more occasions (recall that the DBN model assumes that the top ranked department is always examined,

000083

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

and that examination then proceeds in order of rank). The probability of actual relevance of department #104 (0.2222) is even lower, since it was clicked on often, but without user satisfaction.

```
------------------------------
Training on 10 search sessions (1 unique queries).
------------------------------
       Trained SDBN click model in 0 secs:
SDBNParamNames.attr
3: {'108': 0.8889, '111': 0.7273, '104': 0.6667, '113': 0.6667, '112': 0.6667}

SDBNParamNames.sat
3: {'108': 0.7778, '111': 0.3333, '104': 0.2222, '113': 0.6667, '112': 0.3333}
```

Figure 4: A screenshot of DBN model output for the 10 example user sessions. The row of numbers below SDBNParamNames.attr is for the attractiveness parameters that measure the probability of perceived relevance. The number 3 to the left of the brackets is the query id, and the department ids along with their probabilities are inside the brackets. SDBNParamNames.sat indicates the satisfaction parameters that measure the probability of actual relevance.

Further insight into the model parameter output can be gained by examining small changes to user click patterns. For the 10 user sessions shown above, three have the same pattern {104, 111, 108}. Removing the click on department #111 from one of these three sessions changes it to {104, 108}, which produces the model output shown in Fig. 5. In this case the output data for all departments other than department #111 doesn't change. The probability of perceived relevance for department #111 changes from 0.7273 to 0.6364, since it received one less click that was followed by a click on another department. The probability of actual relevance for department #111 increases, however, from 0.3333 to 0.3750. For user session {104, 111, 108}, it is inferred that #111 is not actually relevant, because it is followed by a click on department #108. No such inference can be made from session {104, 108}, because according to the model, actual relevance cannot be assessed without clicking on the department, meaning that in this case the model has accumulated less evidence that # 111 is not relevant.

Note that while only one unique query was shown in this example, the model is equipped to handle whatever number of unique queries exist in the user logs, where the example output shown in Figs. 4 and 5 is provided for each unique query id.

000084

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

```
------------------------------
Training on 10 search sessions (1 unique queries).
------------------------------
       Trained SDBN click model in 0 secs:
SDBNParamNames.attr
3: {'108': 0.8889, '111': 0.6364, '104': 0.6667, '113': 0.6667, '112': 0.6667}

SDBNParamNames.sat
3: {'108': 0.7778, '111': 0.3750, '104': 0.2222, '113': 0.6667, '112': 0.3333}
```

Figure 5: Model output for the 10 example user sessions with a change from one user session from {104, 111, 108} to {104, 108}.

6 System Integration: Relevance Adjustments

This Section describes how DBN analysis results will be used to adjust the recommendation scores output by the existing Department Recommendation AI system. All of the calculations described here, and in this document in general, happen behind the scenes and therefore integrating the DBN model into the existing recommendation framework won't require any change in the user interface.

6.1 Relevance Adjustment Calculation

Initially, when there are no records in the system logs for a given query, the departmental ranks must depend exclusively on the relevance scores produced by the Department Recommendation AI system (denoted as ρ_{dq}). As user clicks on query results accumulate, the relevance scores generated by the DBN model (denoted as s_{dq}) become more statistically reliable. As such, the adjustments made to the relevance scores for a query q are weighted more heavily to DBN analysis as the number of sessions n_q increases. To achieve this we define an overall relevance score for a query department pair as

$$\Gamma_{dq}(n_q) = w(n_q)\rho_{dq} + (1 - w(n_q))s_{dq} \tag{13}$$

where $w(n_q)$ is the weighting function

$$w(n_q) = e^{-n_q/\tau} \tag{14}$$

and τ is the decay constant, which is set to $\tau = 15$. Applying the weighting function (shown graphically in Fig. 6) allows the influence on Γ_{dq} to change smoothly as user logs for a given query accumulate. As n_q increases, the relevance from the DBN model has increasing influence on the overall relevance score.

000085

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

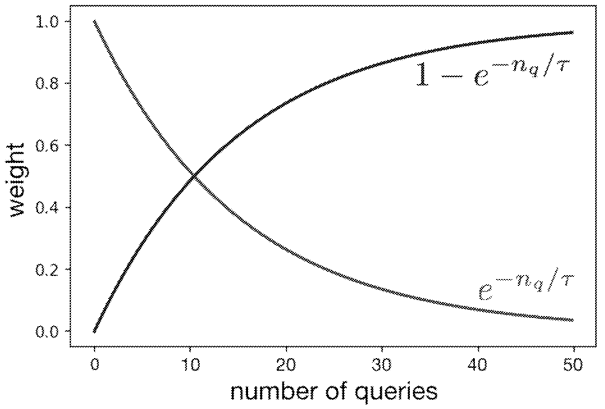


Figure 6: The weighting function as a function of the number of queries. The weight of the AI system recommendation is shown in the red curve, and the corresponding weight for the DBN model recommendation is shown in blue. The decay constant for these curves is $\tau = 15$.

6.2 Example Adjustment

As user queries accumulate, this re-weighting scheme will have the most immediate impact on departments with similar relevance scores. This is because the initial contribution from s_{dq} for small n_q is relatively small, and therefore will only be able to rearrange department rankings that had initial relevance scores that were close to one another.

For example, take two departments with ids #1 and #2 that were returned for a with query id = 5. The departments have $\rho_{1,5} = 0.25$ and $\rho_{2,5} = 0.23$, and after three user interactions with the system ($n_5 = 3$) DBN analysis returns $s_{1,5} = 0.7$ and $s_{2,5} = 0.9$. The weight is therefore $e^{-3/15} = 0.82$, and substitution into Eq. 13 gives $\Gamma_{1,5} = 0.33$ and $\Gamma_{2,5} = 0.35$. The adjusted rankings now indicate that department #2 should be ranked higher than #1. For values of ρ_{dq} that are either very high or low relative to the the other departments, more substantial evidence from DBN analysis of system logs (higher n_q and difference in s_{dq} values) is needed for producing Γ_{dq} that would result in a change in department rank. While $\tau = 15$ is a reasonable for these calculations, we note that system logs can be used in the future to optimize this parameter and other aspects of DBN model integration.

000086

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

7 Conclusions

This document outlines the methodology developed for the Relevance Feedback and System Logs proof of concepts. As it currently stands, the Department Recommendation AI system predicts the relevance of government departments to a given search query. Over time, however, as users interact with the recommendation system, data will accumulate that can be used to infer the actual relevance of departments. Actual relevance can be calculated from data accumulated in system logs using click models. The specific model that fits well within the already existing ATIP recommendation framework is the DBN click model, which is one of the only models that distinguished between perceived and actual relevance. The output from this model can be easily integrated with the current recommendation system to refine and improve the search results provided by the system, which will ultimately lead to increased likelihood of ATIP requests being submitted to the correct department.

000087

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

References

[1] *Department Recommendation: Access To Information and Privacy (ATIP)).* Lixar report, 2018.

[2] Nick Craswell, Onno Zoeter, Michael Taylor, and Bill Ramsey. *An experimental comparison of click position-bias models.* In WSDM, ACM Press, 2008.

[3] Bin Cao, Dou Shen, Kuansen Wang, Qiang Yang. *Clickthrough Log Analysis by Collaborative Ranking.* Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI-10), 224, 2010.

[4] Aleksandr Chuklin, Pavel Serdyukov, and Maarten de Rijke. *Using intent information to model user behavior in diversified search.* In ECIR, 2013.

[5] Olivier Chapelle and Ya Zhang. *A Dynamic Bayesian Network click model for web search ranking.* In WWW, ACM Press, 2009.

[6] Thorsten Joachims, Laura Granka, Bing Pan, Helene Hembrooke, and Geri Gay. *Accurately interpreting clickthrough data as implicit feedback.* In SIGIR, ACM Press, 2005.

[7] Georges E. Dupret and Benjamin Piwowarski. *A user browsing model to predict search engine click data from past observations.* In SIGIR, ACM Press, 2008.

000088

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Lixar deliverables for Feedback and Expedia Style Feedback

Tasks and Deliverables

ATIP Expedia Style Feedback

To provide a means to increase overall user confidence in the ATIP system by delivering timely and impactful feedback alongside the recommendations.

Increasing User Confidence - As per our discussions, and an ongoing focus within the industry, we need to provide a means that encourages adoption and trust within a Machine Learning solution. After further expansion on our initial ideas, we feel that there are two connected means of providing information back to a user to help gain trust and understanding.

Similar Queries - Like many large search portals, where we see suggestions and similarities in user behaviour and search patterns, we can leverage our Search Index and Model solution to provide a list of similar recent queries and phrases seen to be similar.
For instance, a "Did you mean?" suggestion could be served to a user based upon similar received queries which yielded results, and not based on the summaries themselves. This could also be integrated with the Sufficient Information threshold as discussed in our prior PoC documentation.
To enable this, we will utilise the API logs and previously submitted queries, along with Natural Language Processing (NLP) techniques to find and cleanse query strings. Finally, once we have a cleansed and searchable phrase set, we will leverage Solr against the API logs to find similar user queries.
This will be served back via an API, and as part of the current suite as a new method.

Other User Behaviour - To help encourage users and develop a trust within their behaviour, and the AI solution, we could develop an additional API call to analyse the current text content entered by the user.
From here, we can analyse this text, apply NLP techniques, and search both summaries and user queries to identify and return user behaviour statistics.
This technique is often leveraged in search portals and e-commerce sites, to encourage interaction and improve user engagement.

Production Log Analysis & User Feedback

Building on the success of the POCs in the previous phase, Lixar is recommending the productionization of both web log analysis and relevance feedback from user click data. The data provides implicit user feedback that can be mined to fine-tune the relevance scores returned by the Department Recommendation AI system. Integrating relevance feedback into the ATIP system would require significant involvement from the IMTD team as we would be reliant on them capturing and submitting data.

Deliverable - The simplest measurable way that users interact with search results is by clicking on them. Probabilistic models of click-based user interactions with search results are therefore called 'click models'. Click models define user behaviour in terms of events that are either *observable* or *hidden*. Whether or not a particular search result is relevant is hidden; users don't directly label the relevance of search results. What isn't hidden are the observable events, in this case the user clicks.

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

So while relevance isn't immediately observable, click models assume that it depends on the sequence of observable clicks. An appropriate model of user behaviour can therefore be used to assess the relevance of search query results.
Relevance adjustment for the ATIP system will be calculated using a Dynamic Bayesian Network click model. This model captures user behaviour by considering four 'events' that take place during user interaction with search results:

- *E* – *Examination* of a department (the user reads the department name and caption).
- *C* – *Clicking* on a department.
- *A* – The *attraction* of the user to a department and its description. This is the perceived relevance of the department.
- *S* – The *satisfaction* of a user department after it has been clicked. This is the actual relevance of the department.

The main advantage of this click model compared to many others is that it distinguishes between perceived relevance and actual relevance. In the case of the ATIP system, the department with actual relevance is the one where an ATIP request is submitted. STATEMENT OF WORK

Over time, as users interact with the recommendation system, data will accumulate for many unique queries, each of which are given a unique query id. To improve model performance, queries that are not strictly equivalent but are in essence the same will be matched. For example, the queries "the Canadian government" and "government of Canada" are for the same topic, but would be assigned different query ids if strict equivalence was required. A query matching framework was therefore constructed that uses the already developed backbone of the pre-processing procedure for the Department Recommendation AI system. In this procedure, each row of the Document Term Matrix (DTM) is a Term Frequency-Inverse Document Frequency (TF-IDF) vector representation of a given ATIP query, where weights are assigned to terms based on their frequency of appearance within individual documents (TF) relative to the corpus as a whole (IDF). Terms that appear uniquely in queries for specific departments therefore have a higher weight than terms that appear in many departments. Using this approach, queries that contain similar terms will have similar vector representations. When the similarity crosses an appropriate threshold, queries are considered to be the same.
The output from the DBN click model can be easily integrated with the current recommendation system to refine and improve the search results provided by the system, which will ultimately lead to increased likelihood of ATIP requests being submitted to the correct department. To combine the relevance scores produced by the Department Recommendation AI system and the click model, an overall relevance score that combines the results from both will be used. Initially, when there are no records in the system logs for a given query, the overall relevance will depend exclusively on the scores produced by the Department Recommendation AI system. As user clicks on query results accumulate, the relevance scores generated by the click model will become more statistically reliable. As such, the adjustments made to the overall relevance scores for a given query will be weighted more heavily to click model analysis as user interactions for a given query accumulate in the system logs.

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Using artificial intelligence

The search functionality provided on uses AI to improve user experience.

The first instance where AI is used is when searching for information that may have already been released in response to another request. The search results are based on information readily available on the Open Government website.

The second instance where AI is used is when helping to identify which institution may have the information pertaining to the request. The search will recommend institutions that are most suitable for the type of request. The data used to make this recommendation comes from the following locations:

- Open Government summaries
- departmental reports
- "scraping" on government websites
- institutions' ATIP web pages
- Government of Canada taxonomies
- unified master data organization schema
- Part III of Departmental Results Reports for the 2016 to 2017 fiscal year

How are we using AI

Ensuring that a web search finds all the correct documents can be a difficult task. The search system leverages machine learning to identify contextual and latent relationships that are more fundamental than keywords. To do this, the search looks at concepts and the relationship between past searches to improve result quality.

The searching system that was developed uses advanced natural language processing and machine learning techniques to enhance searches across multiple sources. This search solution will include websites, forums or anything that is publicly accessible. By going beyond simple word similarity and instead "understanding" the meaning of search terms, this solution can compare a user's search needs to the corpus of documents in near real time, returning all relevant documents, or components of documents, that relate to a given search query or comparable document.

Synonyms, abbreviations and typos often mean that key documents go overlooked. By using advanced machine learning and natural language processing, the algorithm built is able to read an entire corpus of documents (such as an enterprise website, a course curriculum and the textbooks and related documents). After reading the documents, the AI search system is able to semantically "understand" the phrases and ideas, more so than in keyword matching.

AI algorithm

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

The following will give more technical information about the algorithm that was used:

- category of algorithm used: natural language processing
- models used: tf-idf (term frequency-inverse document frequency) and cosine similarity models

Tf-idf Improvements

Tf-idf is a method to score how related two pieces of text are. The input text from the user will be matched against all documents previously released under ATIP. Public documents that have the highest scores will be suggested to the user.

The tf-idf algorithm begins by counting the number of words in the request that are also present in each public document. This count is then divided by how common each of the matched words are. This division reduces noise and accounts for the fact that common words such as "Canada" are likely to match many documents, regardless of the ATIP request and, therefore, the fact that match isn't as important. It is more valuable to know that a less common word is found in both the user submission and the publicly available document.

At its core, tf-idf is a word-matching algorithm. Similar words (not exact) in both the query and the documents will not register as a match with tf-idf alone. In fact, tf-idf is what powers many off-the-shelf search platforms, including Apache Solr, which ATIP has suggested does not return relevant results. Therefore, a number of improvements to the tf-idf algorithm had to be done.

Stemming

The most common way of improving tf-idf is to use a technique called stemming. Stemming is the process of simplifying a word to its "stem" or root. For example, the root word of "stemming" is "stem." If we reduce all words to their base and then look for matches, we will count two words such as "fishing" and "fisher" as matching. This technique works similarly in English and French.

Stop words

As we move through the content to reduce words to their stem, we can also remove stop words. A stop word is a common word that does not contribute meaning to the phase. For example, if we removed "the" and "a" from any sentence, we can still infer the general meaning. Removing stop words improves the speed of our algorithm and reduces false matches.

Word embeddings

Stemming is useful when two words share the same root. But often there are words that are practically the same but do not share a common root. For example, "Access to Information and Privacy" and "ATIP" have the exact same meaning but share no words in common. In order for tf-idf to register matches for similar words, we need a way of measuring the similarity, or

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

distance, between any two words. For example, "kids" and "children" should be close together, and "sheep" and "lion" should be far apart. In order to measure the distance between words, we can use a method called word embedding.

Word embedding is a tool that converts a word to a vector. Typically, this vector has hundreds of dimensions. We tend to use word embeddings that have 100 to 300 dimensions. Despite having a high number of dimensions, we can calculate the distance between any two words the same way we calculate distance in a smaller number of dimensions.

To combine tf-idf and embeddings, we convert every word to a vector by way of an embedding. We then measure the distance between every word in the activity (source) and every word in the snippet of content (target). Words that are very close together are given a score close to 1 (or exactly 1 if they're the same word), and words that are very far apart are given a score of 0. In this way, a word will be considered a match if the meaning of the word is similar.

Dimensionality reduction

A word embedding converts a single word to a vector of hundreds of numbers. This is done for all words in all publicly available government documents. Ultimately this generates a tremendous amount of data, and this data must be searched and analyzed with every ATIP request. We can reduce the amount of computation (and therefore increase search performance) by using an algorithm called singular value decomposition (SVD). In short, SVD can be used to compress the information in each document (and total data generated) while still retaining the information and search accuracy.

Share this page

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Lauzon, Dominique (he/him, il)

From: Annett, Megan
Sent: November 24, 2023 1:27 PM
To: Trew, Debbie; Herbert, Wesley
Cc: Thompson, Ross
Subject: RE: Recommendation Engine

UNCLASSIFIED / NON CLASSIFIÉ

Thanks Debbie, but you don't need to send it for today.

I think it would be best for you and Wes to review what you find and discover through the testing of the tool to then come up with a comprehensive list of defined requirements that you want moving forward, not just for Ross but also the IMTD team.

Thanks,
megan

From: Trew, Debbie
Sent: Friday, November 24, 2023 1:22 PM
To: Annett, Megan ; Herbert, Wesley
Cc: Thompson, Ross
Subject: RE: Recommendation Engine

UNCLASSIFIED / NON CLASSIFIÉ

Hello
I am still looking around. I have one document I can share, but I really don't know if will help. I will send you what I have by end of day.

Debbie Trew
Senior Project Officer
Open Government, Office of the Chief Information Officer
Treasury Board of Canada Secretariat / Government of Canada
Debbie.Trew@tbs-sct.gc.ca / 613-219-9626

Responsable principal de projet
Division du gouvernement ouvert, Bureau du dirigeant principal de l'information
Secrétariat du Conseil du Trésor du Canada / Gouvernement du Canada
Debbie.Trew@tbs-sct.gc.ca / 613-219-9626

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

From: Annett, Megan <Megan.Annett@tbs-sct.gc.ca>
Sent: Friday, November 24, 2023 1:18 PM
To: Trew, Debbie <Debbie.Trew@tbs-sct.gc.ca>; Herbert, Wesley <Wesley.Herbert@tbs-sct.gc.ca>
Cc: Thompson, Ross <Ross.Thompson@tbs-sct.gc.ca>
Subject: RE: Recommendation Engine

UNCLASSIFIED / NON CLASSIFIÉ

Hi Debbie and Wes,

Checking if there's been any headway with finding previous and/or capturing business requirements, or when you think it might available to share?

Thanks,
megan

From: Trew, Debbie <Debbie.Trew@tbs-sct.gc.ca>
Sent: Wednesday, November 8, 2023 10:20 AM
To: Annett, Megan <Megan.Annett@tbs-sct.gc.ca>; Herbert, Wesley <Wesley.Herbert@tbs-sct.gc.ca>
Cc: Thompson, Ross <Ross.Thompson@tbs-sct.gc.ca>
Subject: RE: Recommendation Engine

UNCLASSIFIED / NON CLASSIFIÉ

Hello
That's great news. Yes please include me in the meeting invite.

Thanks

Debbie Trew
Senior Project Officer
Open Government, Office of the Chief Information Officer
Treasury Board of Canada Secretariat / Government of Canada
Debbie.Trew@tbs-sct.gc.ca / 613-219-9626

Responsable principal de projet
Division du gouvernement overt, Bureau du dirigeant principal de l'information
Secrétariat du Conseil du Trésor du Canada / Gouvernement du Canada
Debbie.Trew@tbs-sct.gc.ca / 613-219-9626

From: Annett, Megan <Megan.Annett@tbs-sct.gc.ca>
Sent: Wednesday, November 8, 2023 10:18 AM
To: Trew, Debbie <Debbie.Trew@tbs-sct.gc.ca>; Herbert, Wesley <Wesley.Herbert@tbs-sct.gc.ca>
Cc: Thompson, Ross <Ross.Thompson@tbs-sct.gc.ca>
Subject: RE: Recommendation Engine

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Hi Debbie,

Ross has deployed the recommendation AI engine (Lixar replacement) in the staging environment. Phillippe Vadnais has access and Ross is meeting with Dominique Tremblay tomorrow to review it. Do you want to be part of that meeting?

Thanks,
megan

From: Trew, Debbie <Debbie.Trew@tbs-sct.gc.ca>
Sent: Wednesday, November 1, 2023 12:28 PM
To: Annett, Megan <Megan.Annett@tbs-sct.gc.ca>; Herbert, Wesley <Wesley.Herbert@tbs-sct.gc.ca>; Thompson, Ross <Ross.Thompson@tbs-sct.gc.ca>
Subject: RE: Recommendation Engine

UNCLASSIFIED / NON CLASSIFIÉ

Hello
When Wes is back in the office I will have a discussion with him to see if there is anything else.

Thanks

Debbie Trew
Senior Project Officer
Open Government, Office of the Chief Information Officer
Treasury Board of Canada Secretariat / Government of Canada
Debbie.Trew@tbs-sct.gc.ca / 613-219-9626

Responsable principal de projet
Division du gouvernement overt, Bureau du dirigeant principal de l'information
Secrétariat du Conseil du Trésor du Canada / Gouvernement du Canada
Debbie.Trew@tbs-sct.gc.ca / 613-219-9626

From: Annett, Megan <Megan.Annett@tbs-sct.gc.ca>
Sent: Wednesday, November 1, 2023 12:06 PM
To: Trew, Debbie <Debbie.Trew@tbs-sct.gc.ca>; Herbert, Wesley <Wesley.Herbert@tbs-sct.gc.ca>; Thompson, Ross <Ross.Thompson@tbs-sct.gc.ca>
Subject: FW: Recommendation Engine

UNCLASSIFIED / NON CLASSIFIÉ

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Hi Debbie,

I believe Ross has a good idea of the recommendation engine from the code review. He's copied here if he wants to review the attached documentation.

Based on the recent talk with Dom, it has been decided that the Expedia piece will not need required, which helps to define the scope of work.

Are there any specific business requirements for the basic recommendation engine that Ross needs to be aware of as he explores solutions?

@Thompson, Ross Is there anything else you need at this stage?

Thanks,
megan

From: Trew, Debbie <Debbie.Trew@tbs-sct.gc.ca>
Sent: Tuesday, October 31, 2023 11:40 AM
To: Annett, Megan <Megan.Annett@tbs-sct.gc.ca>
Cc: Herbert, Wesley <Wesley.Herbert@tbs-sct.gc.ca>
Subject: Recommendation Engine

UNCLASSIFIED / NON CLASSIFIÉ

Hello
I don't have a lot of information on the recommendation engine, but I did find a copy of the TA that explains in some details what Lixar was planning to do. Let me know if this helps at all? If not I will keep looking around to see what I can find.

Debbie Trew
Senior Project Officer
Open Government, Office of the Chief Information Officer
Treasury Board of Canada Secretariat / Government of Canada
Debbie.Trew@tbs-sct.gc.ca / 613-219-9626

Responsable principal de projet
Division du gouvernement overt, Bureau du dirigeant principal de l'information
Secrétariat du Conseil du Trésor du Canada / Gouvernement du Canada
Debbie.Trew@tbs-sct.gc.ca / 613-219-9626

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Lauzon, Dominique (he/him, il)

From: Annett, Megan
Sent: November 24, 2023 2:45 PM
To: Trew, Debbie; Thompson, Ross
Cc: Herbert, Wesley
Subject: RE: Information regarding Recommendation Engine

UNCLASSIFIED / NON CLASSIFIÉ

Thanks Debbie.

I'm not sure what the expectation is for us to do with the attached information. Since the Lixar and OG Portal Recommendation Engine solutions are not apples for apples, what was done in the Lixar product is not necessarily an option for OG side to implement.

The best next step would be for you and Wes to define a high-level list of business requirements/rules for what is needed now. From those, Ross and IMTD can determine if there are technical options that are feasible with the new engine.

Thanks,
megan

From: Trew, Debbie
Sent: Friday, November 24, 2023 2:11 PM
To: Thompson, Ross ; Annett, Megan
Cc: Herbert, Wesley
Subject: Information regarding Recommendation Engine

UNCLASSIFIED / NON CLASSIFIÉ

Hello
Please find noted below and attached information that Wes and I were able to explain the Recommendation Engine. I will continue to look.

==========
==========
This is what Lixar sent us……..

tf-idf - https://en.wikipedia.org/wiki/Tf%E2%80%93idf

cosine similarity - https://en.wikipedia.org/wiki/Cosine_similarity

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

==========

Lixar, we can provide the following preferred schema for all Search and Click Logs.

Preferred Formats: *CSV, JSON, TXT*
Schema:
session_id *Unique identifier for the user session*
datetime *Timestamp for the log*
action_type *"Q" for Query, "C" for Click*
search_term *Query search term used*
click_id *ID of a selected department (applicable when "Click")*
department_ids *An Array of Department UMD IDs; returned by the Lixar API for searches, ordered by Rank (1,2,3, etc)*

I hope this helps, if you would like us to jump on a call to talk around this more, please let us know.

Therefore, we will look to begin development against this schema, however, if this proves difficult or you require an alternative, please let us know as soon as possible.

==========

Sites and files crawled for the Recommendation Engine -
) the summaries files
2) departmental reports --> departments listed on this wesbite https://www.canada.ca/en/treasury-board-secretariat/services/departmental-performance-reports/list-institutions-2016-17-departmental-results-reports.html
3) combing through various gov websites (exact list of url's would take some time to dig out)
4) Institution ATIP Pages.xlsx --> a file shared by the client at the beginning of atip phase 2.
canada.ca
2016-17 Part III - Departmental Results Reports (DRR) - Canada.ca
TBS is providing this listing to make it easier for you to access the Departmental Performance Reports on various government institution websites. Please note that some links might not work until institutions post their reports on their websites.

Here is an example of what the Lixar API returns to us:

{

"StatResults":[

{

"Term":"tax",

"Stats":{

"Last30Days":"0",

"Last90Days":"1",

"MostCommonDepartment":{

"Name":"Canada Revenue Agency | Agence du revenu du Canada",

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

"UmdNumber":"72"

}

}

},

{

"Term":"money",

"Stats":null

},

{

"Term":"canada",

"Stats":{

"Last30Days":"0",

"Last90Days":"1",

"MostCommonDepartment":{

"Name":"Treasury Board of Canada Secretariat | SecrÃ©tariat du Conseil du TrÃ©sor du Canada",

"UmdNumber":"117"

}

}

}

]

}

Debbie Trew
Senior Project Officer
Open Government, Office of the Chief Information Officer
Treasury Board of Canada Secretariat / Government of Canada
Debbie.Trew@tbs-sct.gc.ca / 613-219-9626

Responsable principal de projet

Document released under the *Access to Information Act* /
Document divulgué en vertu de la *Loi sur l'accès à l'information*

Division du gouvernement overt, Bureau du dirigeant principal de l'information
Secrétariat du Conseil du Trésor du Canada / Gouvernement du Canada
Debbie.Trew@tbs-sct.gc.ca / 613-219-9626